

An estimation for the lengths of reduction sequences of the $\lambda\mu\rho\theta$ -calculus

Péter Battyányi

Department of Computer Science, Faculty of Informatics,
University of Debrecen,
Kassai út 26, 4028 Debrecen, Hungary
email : battyanyi.peter@inf.unideb.hu

Karim Nour

LAMA - Équipe LIMD
Université Savoie Mont Blanc, CNRS, LAMA,
F-73000 Chambéry, France
email : karim.nour@univ-smb.fr

Abstract

Since it was realized that the Curry-Howard isomorphism can be extended to the case of classical logic as well, several calculi has appeared as candidates for the encodings of proofs in classical logic. One of them was the $\lambda\mu$ -calculus of Parigot ([19]). In this paper, following the reasoning of Xi presented for the λ -calculus ([27]), we give an upper bound for the lengths of reductions in the $\lambda\mu$ -calculus extended with two more simplification rules: the ρ - and θ -rules.

1 Introduction

1.1 The Curry-Howard isomorphism for classical logic

In the early nineties it was realised that the Curry-Howard isomorphism can be extended to the case of classical logic ([13] and [11]). Since then several calculi have appeared aiming to give an encoding of proofs formulated either in classical natural deduction or in classical sequent calculus ([2], [4], [19], [23]).

One of them was the $\lambda\mu$ -calculus presented by Parigot in [22], which stands very close in nature to the λ -calculus itself. The $\lambda\mu$ -calculus was originated from the notion of Free Deduction ([22]). In order to resolve certain determinism in intuitionistic natural deduction, he uses new kinds of variables, the μ -variables, not active at the moment, but the current continuation can be passed over to them. Eliminating cuts with these new formulas leads to the introduction of a new reduction rule the so-called the μ -rule. The result is a calculus, the $\lambda\mu$ -calculus ([19]), which is in relation with classical natural deduction. In addition, more simplification rules, for example the ρ - and θ -rules, are defined by Parigot ([19] and [20]). The reason for introducing the simplification rules ρ - and θ , together with the symmetric counterpart of the μ -rule, which is the μ' -rule ([20]), was the following. In the typed λ -calculus we are able to define integers by Church's numerals and other data types in the usual manner ([16]). For the Church numerals and the data types the unicity of representation of data holds. This means that, if we talk about the Church numerals only, every term of type N , where N is a type of a Church numeral, is β -equal to a Church numeral. This is no more true for the $\lambda\mu$ -calculus: we can find normal terms of type N that are not church numerals. The problem is resolved by introducing a symmetric equivalent of the μ -rule and some more reduction rules, namely the ρ - and the θ -rules ([20]). We should remark that the price for adding some more rules to the β - and μ -rules was the absence of usual proof theoretical properties, like confluence, for the symmetric $\lambda\mu$ -calculus. Even

strong normalization is lost, when we consider the symmetric $\lambda\mu$ -calculus together with the ρ -rule ([3]). Parigot has showed in [21] that the $\lambda\mu$ -calculus, i.e. when we consider the β - and the μ -rules only, is strongly normalizing: he gave a proof of the result with the help of the Tait-Girard reducibility method [26]. While an arithmetical proof of the same result was presented by David and Nour in [6].

1.2 Outline of the present work

In this paper, based on the argument of Xi ([27]) for the λ -calculus, we present an upper bound for the lengths of reductions in the $\lambda\mu\rho\theta$ -calculus in terms of the complexity and the rank of the term M , defining the rank of M as the maximum of the lengths of the types of the redexes in M . We base our treatment on [3]. The presentation follows somewhat that in [27]. First we prove a standardization theorem for the $\lambda\mu$ -calculus with the additional claim that, in the case of the $\lambda\mu$ I-calculus the length of a reduction sequence is majorized by that of its standardization, which is, in turn, bounded from above by a certain measure defined in the paper. In addition, we show that, if M is a $\lambda\mu$ I-term, then a standard reduction sequence leading to the normal form of M is the leftmost reduction sequence, which is necessarily unique. Thus it makes no difference how we find the standard reduction sequence leading to the normal form of M : it is by all means the longest reduction sequence normalizing M . In Section 4, first we find an appropriate reduction sequence for $\lambda\mu$ I-terms the measure of which can be bounded from above by a super-exponential number theoretic function. Hence, our strategy for a general term M is to define a translation $[M]_k$ of M into the $\lambda\mu$ I-calculus, which depends on the rank k of M and is such that the longest reduction sequence of M is not longer than the longest reduction sequence of $[M]_k$. This is achieved in the rest of the section. Since we have already obtained a bound for the reduction sequences of $[M]_k$, we also have one for those of M . Moreover, we obtain the same bound as that of Xi [27], except for that we calculate the ranks of terms by taking into consideration not only the β -redexes but the μ -, ρ - and θ -redexes also.

1.3 Some difficulties and our proposed solutions

The new redexes impose additional difficulties: first of all a notion of a standard reduction sequence had to be created which is vastly different in appearance from that for the λ -calculus. Naturally, the standard reduction sequences of purely λ -terms remain the same, but, with the presence of new redexes overlaps can occur: performing a θ -redex can vanish a μ -redex and executing a θ -redex can make a ρ -redex disappear and vice versa. Our definition of a standard reduction sequence excludes these overlaps: a redex which can be a θ - and a μ -redex simultaneously will always be considered as a μ -redex. Likewise, a θ - or a ρ -redex is not allowed to be performed when a ρ - or a θ -redex would be destroyed in the same time. We show that our choice is appropriate: we can always find standard reduction sequences respecting these stipulations. Due to the presence of other reduction rules, finding the bound and proving that the lengths of the reduction sequences obey that bound is more difficult even in the case of $\lambda\mu$ I-terms. Instead of giving bounds for the lengths of arbitrary normalizing reduction sequences and sizes of normal forms, as Xi is able to do for the λ -calculus ([28]), we estimate these features only for special kinds of reduction sequences which we call good k -normalization sequences. We calculate the bounds for the general case by assigning a $\lambda\mu$ I-term $[M]_k$ with a certain k to the $\lambda\mu$ -term M such that the longest reduction paths of $[M]_k$ are at least as long those of M . Concerning the general case, we apply modified transformation than that of Xi ([27]), the original idea of Xi contains a slight impreciseness: when M is a λ -abstraction, we have to apply case distinction deciding whether M is the

left hand side of a β -redex or not. Otherwise the length of the assigned $\lambda\mu$ I-term could not be estimated correctly. Finally, in the Appendix we present a formal treatment of the notions of reduction sequences, residuals, involvement of redexes, which will be used in an informal way throughout the paper.

2 The $\lambda\mu\rho\theta$ -calculus

2.1 The syntax of calculus

The $\lambda\mu\rho\theta$ -calculus was introduced by Parigot in [19]. Instead of his original calculus, we use a modified version owing to de Groote [12]: we apply the term formation rules in a more flexible form, we do not assume that a μ -prefix must be followed by a μ -variable. In what follows, we give the appropriate definitions.

Definition 2.1 (Terms) 1. *There are two kinds of variables : the set of λ -variables $\mathcal{V} = \{x, y, z, \dots\}$ and the set of μ -variables $\mathcal{W} = \{\alpha, \beta, \gamma, \dots\}$. The set of terms is denoted by \mathcal{T} and the term formation rules are:*

$$\mathcal{T} := \mathcal{V} \mid \lambda x \mathcal{T} \mid \mu \mathcal{W} \mathcal{T} \mid (\mathcal{W} \mathcal{T}) \mid (\mathcal{T} \mathcal{T}).$$

2. *The complexity of a term is defined as follows:*

- $comp(x) = 1$,
- $comp((\alpha M)) = comp(\lambda x M) = comp(\mu \alpha M) = comp(M) + 1$,
- $comp((M N)) = comp(M) + comp(N)$.

3. *As usual we denote by $Fv(M)$ the set of free variables of term M .*

4. *Let M and N be terms. We write $N \leq M$ if N is a subterm of M and $N < M$, if $N \leq M$ and $N \neq M$.*

The calculus examined by us is the simply typed one.

Definition 2.2 (Type system) 1. *The types are built from atomic formulas (or, in other words, atomic types) and the constant symbol \perp with the connector \rightarrow . As usual for every type A , $\neg A$ is an abbreviation for $A \rightarrow \perp$.*

2. *The length of a type A (denoted by $lh(A)$) is defined as the number of arrows of A .*

3. *In the definition below Γ denotes a (possibly empty) context, that is, a finite set of declarations of the form $x : A$ (resp. $\alpha : \neg A$) for a λ -variable x (resp. a μ -variable α) and type A such that a λ -variable x (resp. a μ -variable α) occurs at most once in an expression $x : A$ (resp. $\alpha : \neg A$) of Γ . The typing rules are:*

$$\begin{array}{c} \overline{\Gamma, x : A \vdash x : A}^{ax} \\[10pt] \frac{\Gamma, x : A \vdash M : B}{\Gamma \vdash \lambda x M : A \rightarrow B} \rightarrow_i \qquad \frac{\Gamma \vdash M : A \rightarrow B \quad \Gamma \vdash N : A}{\Gamma \vdash (M N) : B} \rightarrow_e \\[10pt] \frac{\Gamma, \alpha : \neg A \vdash M : A}{\Gamma, \alpha : \neg A \vdash (\alpha M) : \perp} \perp \qquad \frac{\Gamma, \alpha : \neg A \vdash M : \perp}{\Gamma \vdash \mu \alpha M : A} \mu \end{array}$$

4. We will say that a term M is typable with A , if there is a set of declarations Γ such that $\Gamma \vdash M : A$ holds.

Definition 2.3 (Reduction rules) 1. We have four kinds of redexes. A β -redex (term of the form $(\lambda x M N)$), a μ -redex (term of the form $(\mu \alpha M N)$), a ρ -redex (term of the form $(\alpha \mu \beta M)$) and a θ -redex (term of the form $\mu \alpha (\alpha M)$ and $\alpha \notin Fv(M)$). We denote by NF the set of normal forms i.e. terms without redex.

2. The reduction rules, aiming to reduce redexes of terms, are as follows:

- The β -reduction rule is $(\lambda x M N) \rightarrow_\beta M[x := N]$
where $M[x := N]$ is obtained from M by replacing every x in M by N .
 - The μ -reduction rule is $(\mu \alpha M N) \rightarrow_\mu \mu \alpha M[\alpha :=_r N]$
where $M[\alpha :=_r N]$ is obtained from M by replacing every subterm in M of the form (αU) by $(\alpha (U N))$.
 - The ρ -reduction is $(\alpha \mu \beta M) \rightarrow_\rho M[\beta := \alpha]$
where $M[\beta := \alpha]$ is obtained by exchanging in M every free occurrence of β for α .
 - The θ -reduction is $\mu \alpha (\alpha M) \rightarrow_\theta M$ provided $\alpha \notin Fv(M)$.
3. Let R be a redex of M . We write $M \rightarrow^R N$ if N is the term M after the reduction of R . If $M = M_1 \rightarrow^{R_1} M_2 \rightarrow^{R_2} \dots \rightarrow^{R_n} M_{n+1} = N$, then $\sigma = [R_1, \dots, R_n]$ denotes this reduction sequence, $n = |\sigma|$ and we write $M \rightarrow^\sigma N$.
4. Let σ, ν be (possibly empty) sequences of reductions. Then $\sigma \# \nu$ denotes their concatenation. Let $\sigma = [R_1, \dots, R_n]$. We denote by $\sigma[x := M]$ (resp. $\sigma[\alpha :=_r M]$) the reduction sequence $[R_1[x := M], \dots, R_n[x := M]]$ (resp. $[R_1[\alpha :=_r M], \dots, R_n[\alpha :=_r M]]$). Moreover, let $\sigma[\alpha := \beta]$ denote the reduction sequence $[R_1[\alpha := \beta], \dots, R_n[\alpha := \beta]]$.
5. As it is customary, by a reduction step we mean the closure of the reduction relation compatible with respect to the term formation rules. In general \rightarrow denotes the compatible closure of a reduction relation, or that of the union of some set of relations, while by \rightarrow^* we mean the reflexive, transitive closure of \rightarrow . Sometimes We write $M \rightarrow^n N$ if M is reduced with n steps of reductions to N .
6. If M is in strongly normalizing i.e. M has no infinite reduction, then, by König's infinity lemma, $\eta(M)$ will denote the length of the longest reduction starting from M .

We present below some theoretical properties of the $\lambda\mu\rho\theta$ -calculus.

Theorem 2.4 (Church-Rosser property) Let M_1, M_2 and M_3 be terms such that $M_1 \rightarrow^* M_2$ and $M_1 \rightarrow^* M_3$. Then there exists an M_4 for which $M_2 \rightarrow^* M_4$ and $M_3 \rightarrow^* M_4$.

A proof of the above assertion can be found in Parigot [19], in Py [24] or in Rozière [25]. In Py [24] the question is expounded to a greater extent together with the results belonging to the theme.

Proposition 2.5 (Type preservation property) Let M, N and A, Γ be such that $\Gamma \vdash M : A$ and $M \rightarrow^* N$. Then $\Gamma \vdash N : A$.

The property can be verified by double induction on the length of the reduction sequence $M \rightarrow N$ and the complexity of M .

Theorem 2.6 (Strong normalization) *If M is a typable term, then M is strongly normalizing i.e. every reduction sequence starting from M is finite.*

There are several proofs of this result in the literature. Consider, for example, Parigot [21], David and Nour [6]. In [14] de Groote proves the strong normalization of the simply typed $\lambda\mu$ -calculus extended with terms of conjunctive and disjunctive types, respectively. He does not consider the ρ - and θ -reduction rules in his calculus.

In this paper we consider only simply typed $\lambda\mu$ -terms, which involves that a μ -variable cannot have but one argument, in addition to this its argument must not be a λ -abstraction.

2.2 Head and leftmost reductions

In order to proceed to the standardization result, we define the notions of head- and leftmost reduction sequences. Both are special cases of standard reduction sequences discussed in the next section.

- Definition 2.7**
1. Let M be a term and \vec{P} a possibly empty sequence of terms M_1, \dots, M_n . We write $(M \vec{P})$ for the term $(\dots((M M_1) M_2) \dots M_n)$ denoted also by $(M M_1 \dots M_n)$.
 2. Let $M = (M_1 \dots M_n) = (M_1 \vec{P})$, with a possibly empty sequence of terms \vec{P} . Then, for $2 \leq i \leq n$, we write $M_i \in \vec{P}$ and we call the M_i ($2 \leq i \leq n$) the components of \vec{P} and the arguments of M_1 .
 3. Let $M = (M_0 M_1 \dots M_n) = (M_0 \vec{P})$. We write $\vec{P} \rightarrow^\sigma \vec{P}'$, when $\sigma = \sigma_1 \# \dots \# \sigma_n$ such that $M_i \rightarrow^{\sigma_i} M'_i$ ($1 \leq i \leq n$).

Lemma 2.8 *Every term of the simply typed $\lambda\mu\rho\theta$ -calculus M can be written uniquely in one of the following forms.*

1. M is a variable,
2. $M = \lambda x M_1$,
3. $M = \mu \alpha M_1$ and M is not a θ -redex,
4. $M = (\alpha M_1)$ and M is not a ρ -redex,
5. $M = (x M_1 \vec{P})$,
6. $M = (\lambda x M_1 M_2 \vec{P})$,
7. $M = (\mu \alpha M_1 M_2 \vec{P})$,
8. $M = (\alpha \mu \beta M_1)$,
9. $M = \mu \alpha (\alpha M_1)$ and $\alpha \notin Fv(M_1)$.

Proof By induction on $comp(M)$. □

Definition 2.9 1. The head-redex of a term M is defined as follows. The numbering of the cases refers to the numbering in Lemma 2.8.

- Case 1: M has no head-redex.
- Cases 2-3-4: The head-redex of M is that of M_1 , if it exists.
- Case 5: M has no head-redex.
- Case 6: The head-redex of M is $(\lambda x M_1 M_2)$.
- Case 7: The head-redex of M is $(\mu \alpha M_1 M_2)$.
- Case 8-9: The head-redex of M is M itself.

If $M = (\mu \alpha M_1 M_2 \vec{P})$, which is case 7, then there can appear a critical pair of redexes provided $\mu \alpha M_1$ is a θ -redex as well. In this situation we always choose the μ -redex $(\mu \alpha M_1 M_2)$ as the head-redex of M .

2. Let $M_1 \rightarrow^{R_1} M_2 \rightarrow^{R_2} \dots \rightarrow^{R_n} M_{n+1}$. Then $\sigma = [R_1, \dots, R_n]$ is a head-reduction sequence, if, for each $1 \leq i \leq n$, R_i is the head-redex of M_i . We denote by $M \rightarrow_{hd} N$ the fact that M reduces to N via a head-reduction sequence.

Definition 2.10 1. The leftmost-redex of a term M , in notation $lr(M)$, is defined as follows.

- $lr(\lambda x M) = lr(M)$,
- $lr(\mu \alpha M) = \mu \alpha M$ if $\mu \alpha M$ is a θ -redex, and $lr(\mu \alpha M) = lr(M)$ otherwise.
- $lr((\alpha M)) = (\alpha M)$ if (αM) is a ρ -redex, and $lr((\alpha M)) = lr(M)$ otherwise,
- $lr((\mu \alpha M_1 M_2 \vec{P})) = (\mu \alpha M_1 M_2)$,
- $lr((\lambda x M_1 M_2 \vec{P})) = (\lambda x M_1 M_2)$,
- $lr((x M_1 M_2 \dots M_n)) = lr(M_i)$ provided $M_i \notin NF$ and $M_j \in NF$ ($1 \leq j \leq i-1$).

2. A reduction sequence $M_1 \rightarrow^{R_1} M_2 \rightarrow^{R_2} \dots \rightarrow^{R_n} M_{n+1}$ is the leftmost-reduction sequence from M_1 to M_{n+1} if R_i is the leftmost-redex of M_i ($1 \leq i \leq n$). We denote by $M \rightarrow_{lrs} N$ the fact that M reduces to N via a leftmost-reduction sequence. Then the reduction sequence itself is denoted by $lrs(M \rightarrow N)$. If $M \rightarrow^\sigma N$ and σ is a leftmost reduction sequence, then σ is unique.

In the literature a special case of the leftmost-reduction strategy is also well-known, this is the notion of head-reduction ([1]). In what follows we compare briefly the two notions of reductions.

Lemma 2.11 Every head-reduction sequence is a leftmost-reduction sequence.

Proof A straightforward induction on the complexity of the term, comparing the various subcases of Definitions 2.9 and 2.10. \square

We give a sketch of the proof that every leftmost-reduction sequence is the concatenation of head-reduction sequences, however. To this end, we first settle what we mean by a term being in head-normal form.

Definition 2.12 *A term M is in head-normal form (in notation $M \in HNF$), if one of the following cases is valid.*

- $M = \lambda x M_1$ and $M_1 \in HNF$,
- $M = (x M_1 \dots M_k)$,
- $M = \mu \alpha M_1$, M is not a θ -redex and $M_1 \in HNF$,
- $M = (\alpha M_1)$, M is not a ρ -redex and $M_1 \in HNF$.

We say that $M' \in HNF$ is a head-normal form of M , if $M \rightarrow_{hd} M'$. Observe that, since the typed $\lambda\mu\rho\theta$ -calculus is strongly normalizing, every term has a unique head-normal form.

Prior to detailing the connection between leftmost reduction and head reduction, we introduce a new notion.

Definition 2.13 *Let $M \in HNF$.*

1. *The core of M , in notation $core(M)$, is defined as follows.*

- *If $M = \lambda x M_1$ or $M = \mu \alpha M_1$ or $M = (\alpha M_1)$, then $core(M) = core(M_1)$.*
- *If $M = x$ or $M = (M_1 M_2)$, then $core(M) = M$.*

Observe that, if $M \in HNF$, $core(M)$ can be obtained from M if we omit the initial λ -, μ -prefixes or μ -variables standing in front of M .

2. *Assume $core(M) = (x \vec{P})$ with a possibly empty \vec{P} . Then we call the components of \vec{P} the components of M , as well.*

Intuitively, a leftmost reduction sequence is a head reduction sequence until the term reaches a head normal form. At this point, the leftmost reduction sequence is the concatenation of the leftmost reduction sequences of the components. This is the content of the lemma below.

Lemma 2.14 *Let $M \rightarrow^\sigma M''$ be a leftmost reduction sequence. Then there exists $M' \in HNF$ and σ' , σ'' such that $M \rightarrow^{\sigma'} M' \rightarrow^{\sigma''} M''$, where σ' is a head reduction sequence and, if $core(M') = (x M'_1 \dots M'_k)$, then $\sigma'' = \nu_1 \# \dots \# \nu_k$ such that $core(M'') = (x M''_1 \dots M''_k)$, $M'_i \rightarrow^{\nu_i} M''_i$ and ν_i are leftmost reduction sequences ($1 \leq i \leq k$).*

Proof By induction on $comp(M)$ taking into account the subcases of Definition 2.10. Let $M \rightarrow_{lrs}^\sigma M''$ and assume $\sigma = [R] \# \hat{\sigma}$. A straightforward observation of the points of Definitions 2.9 and 2.10 gives that, if $M'' \notin HNF$, then R is the head redex of M . Hence we may assume $M \rightarrow^{\sigma''} M''$, where $M \in HNF$. Then, by induction on $cxy(M)$, we obtain that we may suppose that $M = x$ or $M = (x M_1 \dots M_n)$. Both suppositions, by Definition 2.10, immediately yield the result. \square

2.3 Other definitions

In the next section we also formulate statements for $\lambda\mu I$ -terms. We define here the notion of a $\lambda\mu I$ -term and a $\lambda\mu I$ -redex, together with some main properties of the $\lambda\mu I$ -calculus.

Definition 2.15 1. The set of $\lambda\mu I$ -terms is defined inductively as follows:

- x is a $\lambda\mu I$ -term,
- $\lambda x M$ is a $\lambda\mu I$ -term provided M is a $\lambda\mu I$ -term and $x \in Fv(M)$,
- $(M N)$ is a $\lambda\mu I$ -term if M and N are $\lambda\mu I$ -terms,
- $\mu\alpha M$ is a $\lambda\mu I$ -term provided M is a $\lambda\mu I$ -term and $\alpha \in Fv(M)$,
- (αM) is a $\lambda\mu I$ -term provided M is a $\lambda\mu I$ -term.

2. If $M = (\lambda x M_1 M_2)$ (resp. $M = (\mu\alpha M_1 M_2)$), where $\lambda x M_1$ (resp. $\mu\alpha M_1$) and M_2 are $\lambda\mu I$ -terms, then M is called a $\lambda\mu I$ -redex.

It is easy to see that if M is a $\lambda\mu I$ -term and $M \rightarrow M'$, then M' is a $\lambda\mu I$ -term, too. Thus, it is clear that this calculus also has the following three properties: Church-Rosser, type preservation and strong normalization.

The next chapter is concerned with a standardization result in the $\lambda\mu\rho\theta$ -calculus. In the sequel, we are going to use the notions of subterms, redexes, reduction sequences, residuals etc. in an intuitive manner. The precise notions can be found in the Appendix, where we build a treatment of these concepts based on term occurrences marked by individual indexes. A reduction sequence $M_1 \rightarrow^{R_1} M_2 \rightarrow^{R_2} \dots \rightarrow^{R_n} M_{n+1}$ is a sequence of terms and redex occurrences, where M_{i+1} is obtained by reducing with R_i in M_i ($1 \leq i \leq n$). In what follows, as an abuse of notation, the sequence will be referred to without explicitly determining the exact occurrences of the redexes in the terms, if they are clear from the context. We give a short account of the intuitive notions for residuals and involvement of redexes. For a detailed treatment we refer the reader to the Appendix.

Definition 2.16 1. Let $M \rightarrow^R M'$ be a reduction step.

- (a) If $R = (\lambda x R_1 R_2) < M$, then R and $\lambda x R_1$ have no residuals, otherwise, if $(\lambda x R_1 R_2) < U \leq M$, we obtain the residual of U by exchanging $(\lambda x R_1 R_2)$ for $R_1[x := R_2]$ in U . When $U \leq R_1$, then we obtain the residual by substituting each occurrence of x by R_2 in U . In the case of $U \leq R_2$, the residual of U is the same, only its position changes in M' : its index will be one of the indices of a former occurrence of x in R_1 . If R and U are disjoint, then the residual of U is U itself.
- (b) The situations are analogous in the cases of the other redexes: if $R = (\mu\alpha R_1 R_2)$, then R has no residual, if $R = (\alpha \mu\beta R_1)$, then R and $\mu\beta R_1$ have no residuals, and, finally, if $R = \mu\alpha(\alpha R_1)$, then R and (αR_1) have no residuals. Besides these afore-said cases, if $R < U \leq M$, then we obtain the residual of U , if we perform the redex R in U . When $U < R$ and U has a residual: if $R = (\mu\alpha R_1 R_2)$ and $U \leq R_1$, we obtain the residual by recursively exchanging every subterm (αP) of U by $(\alpha (P R_2))$. If $U \leq R_2$, then the residual is the same only its index changes in M' . Otherwise, for $U \leq R_1$ and $R = (\alpha \mu\beta R_1)$, the residual is $U[\beta := \alpha]$. If $U \leq R_1$ and $R = \mu\alpha(\alpha R_1)$, then the residual is U just the index is modified in M' .

2. *Residuals of terms with respect to reduction sequences are defined in a recursive way: we obtain the residuals with respect to a reduction sequence if we compute the residuals of the residuals with respect to subsequences of the reduction sequence.*
3. *Let σ be the reduction sequence $M_1 \rightarrow^{R_1} M_2 \rightarrow^{R_2} \dots \rightarrow^{R_n} M_{n+1}$, assume $R \leq M_1$ is a redex. Then we say that R is involved in σ , if there is an $1 \leq i \leq n$ such that $R = R_i$ and R_i is a residual of R with respect to $M_1 \rightarrow^{R_1} \dots \rightarrow^{R_{i-1}} M_i$.*

In what follows, most of the proofs will follow an induction on a lexicographically ordered pair of integers. Ordering of tuples is understood in the usual lexicographic manner: $(n, m) \leq (n', m')$ iff either $n < n'$ or $n = n'$ and $m \leq m'$.

3 Standardization for the $\lambda\mu\rho\theta$ -calculus

In the present subsection we inspect some assertions concerning estimations for the lengths of standard reduction sequences in the $\lambda\mu\rho\theta$ -calculus. We are concerned with the lengths of standard reduction sequences of the $\lambda\mu$ -calculus together with the ρ - and θ -rules and, in the end of section, we formulate some statements concerning the lengths of standard reductions of the $\lambda\mu$ I-calculus. Many of the proofs are adaptations of the ones related to the simply typed λ -calculus in [27]. The result itself, however, is not a simple generalization of Xi's method. In the presence of μ -, ρ - and θ -reductions there may be overlapping redexes. We define the notion of a standard reduction sequence such that every standard reduction sequence should obey the following properties. When a redex, which is simultaneously a μ - and a θ -redex, involved in a standard reduction sequence, we stipulate that the redex should be understood only as a μ -redex. Likewise, when a θ -redex would destroy a μ -redex we consider reducing the μ -redex, only, and when a μ -redex would make a θ -redex disappear, we forbid the μ -redex. These raise additional issues in the estimation of the lengths of standard reduction sequences: we must take into account the numbers of arguments of θ -redexes which are simultaneously μ -redexes, as well. These are reflected in the definition of a standard reduction sequences and in the measures for the redexes presented in Definition 3.18. We show that in case of overlapping redexes our choice is appropriate: we can majorize by length every reduction sequence by a standard reduction sequence of Definition 3.1. This is the main result of this section: the standardization theorem for the $\lambda\mu\rho\theta$ -calculus.

We should remark that the widely known and intuitive definition demands of a standard reduction sequence that no redex is a residual of a redex which lied to the left of some earlier element of the sequence ([1]). Instead of this, we use a definition of a standard reduction sequence similar to the one applied in [5], which enables us to prove properties concerning standard reduction sequences by induction on the complexity of terms.

3.1 Standard reduction sequences in the $\lambda\mu\rho\theta$ -calculus

In this subsection we define the notion of a standard $\beta\mu\rho\theta$ -reduction sequence and present some elementary lemmas concerning their properties.

Definition 3.1 *A reduction sequence $M \rightarrow^\sigma N$ is standard if one of the following cases holds.*

1. $M = \lambda x M_1$, $N = \lambda x N_1$, $M_1 \rightarrow^\sigma N_1$ and σ is standard.
2. If $M = \mu\alpha M_1$, let $M \rightarrow^\sigma N$ be $M = P_1 \rightarrow P_2 \rightarrow \dots \rightarrow P_{k+1} = N$.

- (a) Either $N = \mu\alpha N_1$, $M_1 \twoheadrightarrow^\sigma N_1$ and σ is standard and none of P_j is a θ -redex ($1 \leq j \leq k+1$),
 - (b) or let $M \twoheadrightarrow^{\sigma'} \mu\alpha(\alpha M') = P_j$ such that P_j is the first term in the sequence which is a θ -redex and σ' is standard and
 - i. either $\mu\alpha(\alpha M') = P_j \rightarrow_\theta M' \twoheadrightarrow^{\sigma''} N$,
 - ii. or $\mu\alpha(\alpha M') = P_j \twoheadrightarrow^{\sigma'''} \mu\alpha(\alpha N')$ such that $M' \twoheadrightarrow^{\sigma'''} N'$,
 where σ'' and σ''' are standard.
3. If $M = (\alpha M_1)$, let $M \twoheadrightarrow^\sigma N$ be $M = P_1 \rightarrow P_2 \rightarrow \dots \rightarrow P_{k+1} = N$.
- (a) Either $N = (\alpha N_1)$, $M_1 \twoheadrightarrow^\sigma N_1$ and σ is standard and none of P_j is a ρ -redex ($1 \leq j \leq k+1$),
 - (b) or let $M \twoheadrightarrow^{\sigma'} (\alpha \mu\beta M') = P_j$ such that P_j is the first term in the sequence which is a ρ -redex and σ' is standard and
 - i. either $(\alpha \mu\beta M') = P_j \rightarrow_\rho M'[\beta := \alpha] \twoheadrightarrow^{\sigma''} N$,
 - ii. or $(\alpha \mu\beta M') = P_j \twoheadrightarrow^{\sigma'''} (\alpha \mu\beta N')$ such that $M' \twoheadrightarrow^{\sigma'''} N'$,
 where σ'' and σ''' are standard.
4. $M = (\lambda x M_1 M_2 \dots M_n)$ and
- (a) either $M \rightarrow_\beta (M_1[x := M_2] \dots M_n) \twoheadrightarrow^{\sigma_1} N$ and σ_1 is standard,
 - (b) or $M \twoheadrightarrow^{\sigma_1} (\lambda x N_1 M_2 \dots M_n) \twoheadrightarrow^{\sigma_2} (\lambda x N_1 N_2 \dots M_n) \twoheadrightarrow^{\sigma_3} \dots \twoheadrightarrow^{\sigma_n} (\lambda x N_1 N_2 \dots N_n) = N$ and σ_i ($1 \leq i \leq n$) are standard.
5. $M = (\mu\alpha M_1 M_2 \dots M_n)$ and
- (a) either $M \rightarrow_\mu (\mu\alpha M_1[\alpha :=_r M_2] \dots M_n) \twoheadrightarrow^{\sigma_1} N$ and σ_1 is standard,
 - (b) or $M \twoheadrightarrow^{\sigma_1} (\mu\alpha N_1 M_2 \dots M_n) \twoheadrightarrow^{\sigma_2} (\mu\alpha N_1 N_2 \dots M_n) \twoheadrightarrow^{\sigma_3} \dots \twoheadrightarrow^{\sigma_n} (\mu\alpha N_1 N_2 \dots N_n) = N$ and σ_i ($1 \leq i \leq n$) are standard.
6. $M = (x M_1 \dots M_n)$, $M = (x M_1 \dots M_n) \twoheadrightarrow^{\sigma_1} (x N_1 \dots M_n) \twoheadrightarrow^{\sigma_2} \dots \twoheadrightarrow^{\sigma_n} (x N_1 N_2 \dots N_n) = N$ and σ_i ($1 \leq i \leq n$) are standard.

In the rest of this paper, we may see a reduction sequence σ as a list of terms of σ or sometimes as the list of the redex occurrences of the reduction sequence. In accordance with this, given a standard reduction sequence $M_1 \rightarrow^{R_1} M_2 \rightarrow^{R_2} \dots \rightarrow^{R_n} M_{n+1}$, we may say that the sequence M_1, \dots, M_{n+1} is standard (the redex occurrences are implicitly understood in M_i), or we may talk about the same thing by just saying that the sequence $\sigma = [R_1, \dots, R_n]$ is standard. In notation: $\sigma \in St$.

We illustrate some of the difficulties in the example below, when we want to assert statements about standard reduction sequences.

Example 3.2 Let $M = (\mu\alpha(\alpha \lambda u(\mu\beta(\beta \lambda yx) (\alpha x))) x)$. Then, if we choose the θ -redex $\mu\beta(\beta \lambda yx)$, we obtain $M = (\mu\alpha(\alpha \lambda u(\mu\beta(\beta \lambda yx) (\alpha x))) x) \rightarrow_\theta (\mu\alpha(\alpha \lambda u(\lambda yx (\alpha x))) x)$, and, since we are not allowed to reduce the redex $(\lambda yx (\alpha x))$, there are no more reductions provided we restrict ourselves to standard ones. On the other hand, if we choose the μ -redex $(\mu\beta(\beta \lambda yx) (\alpha x))$, then $M = (\mu\alpha(\alpha \lambda u(\mu\beta(\beta \lambda yx) (\alpha x))) x) \rightarrow_\mu (\mu\alpha(\alpha \lambda u(\mu\beta(\beta (\lambda yx (\alpha x)))) x) \rightarrow_\beta (\mu\alpha(\alpha \lambda u(\mu\beta(\beta x))) x) \rightarrow_\theta (\lambda u(\mu\beta(\beta x)) x) \rightarrow_\beta \mu\beta(\beta x) \rightarrow_\theta x$ is standard.

It is obvious that standard reduction sequences are not necessarily left to right, in contrast with the situation in the λ -calculus, and this is the case even in the $\lambda\mu$ -calculus without the auxiliary rules ρ and θ .

Our aim in this section is to obtain a standardization theorem for the $\lambda\mu\rho\theta$ -calculus, together with an upper bound of the lengths of the standard reduction sequences. To this end, we state and prove some auxiliary propositions first concerning the behaviour of standard reduction sequences and then we present some lemmas providing upper bounds for the lengths of reduction sequences of terms obtained as substitutions.

We state our first theorem saying that left-most reduction sequences are special cases of standard reduction sequences.

Theorem 3.3 *Every leftmost reduction sequence is standard.*

Proof Immediate from Definitions 2.10 and 3.1. \square

The following lemma states that a reduction sequence, which consists of a head reduction sequence followed by a standard reduction sequence is itself standard.

Lemma 3.4 *Let $M \rightarrow^{\sigma'} M' \rightarrow^{\sigma''} M''$ such that σ' is a head-reduction sequence and σ'' is standard. Then $\sigma = \sigma' \# \sigma''$ is standard.*

Proof Let $\sigma' = [R] \# \nu$. We prove the result by induction on $(|\sigma'|, \text{comp}(M))$, taking into account the various points of Definition 2.9. Assume $|\nu| = 0$. We deal only two cases.

1. $M = (\alpha M_1)$.
 - (a) Assume M is a ρ -redex, then $M_1 = \mu\beta M_2$ and $M \rightarrow_\rho M_2[\beta := \alpha] \rightarrow^{\sigma''} M''$ is standard by point 3 of Definition 3.1.
 - (b) If $R \leq M_1$, then the induction hypothesis applies.
2. $M = (\mu\alpha M_1 M_2 \dots M_n)$. In this case the head redex of M is $(\mu\alpha M_1 M_2)$. Thus $M \rightarrow_\mu (\mu\alpha M_1[\alpha :=_r M_2] \dots M_n) \rightarrow^{\sigma''} M''$ and σ is standard by point 5 of Definition 3.1.

The cases when $|\nu| > 0$ follow from the induction hypothesis. \square

The lemma below is a technical one, it will be useful for verifying standardness of substitutions in the next subsection.

Lemma 3.5 1. *Let $M = (M_1 M_2) \rightarrow^\sigma (\mu\alpha M_3 M_2) \rightarrow_\mu (\mu\alpha M_3[\alpha := M_2]) \rightarrow^\nu N$ such that $\sigma, \nu \in St$ and suppose $\mu\alpha M_3$ is the first term of the reduction sequence $M_1 \rightarrow^\sigma \mu\alpha M_3$ of the form $\mu\beta P$. Then $\xi = \sigma \# [(\mu\alpha M_3 M_2)] \# \nu$ is standard.*

2. *Let $M = (M_1 M_2) \rightarrow^\sigma (\lambda x M_3 M_2) \rightarrow_\mu M_3[x := M_2] \rightarrow^\nu N$ such that $\sigma, \nu \in St$ and suppose $\lambda x M_3$ is the first term of the reduction sequence $M_1 \rightarrow^\sigma \lambda x M_3$ of the form $\lambda y P$. Then $\xi = \sigma \# [(\lambda x M_3 M_2)] \# \nu$ is standard.*

Proof We deal only with case 1. We examine some of the interesting cases. The proof goes by induction on $(|\sigma|, \text{ctxty}(M))$ taking into account the various points of Definition 3.1. If σ is standard by virtue of point 2 of Definition 3.1, that is, $M = \mu\alpha M_1$, the only possibility is when $R = \mu\alpha(\alpha R_1)$ is the first θ -redex in the sequence. But then R is in fact M , and $[R] \# \nu$ is standard by definition. Let σ be standard by reason of point 5 of Definition 3.1. By assumption, the only possibility is $M = (\mu\alpha P_1 P_2 \dots P_n)$ and $\sigma = [(\mu\alpha P_1 P_2)] \# \sigma'$. The induction hypothesis can be applied to σ' and $M' = (\mu\alpha P_1[\alpha :=_r P_2] \dots P_n)$. \square

The following lemma gives us some information on the form of a term which is a redex obtained by a standard reduction sequence.

Lemma 3.6 *Let $M \twoheadrightarrow^\sigma M'$ such that σ is standard. Assume the head-redex $hd(M)$ of M , if it exists, is not involved in σ . Then the following statements are true.*

1. If $M = \lambda x M_1$, then $M' = \lambda x M'_1$, where $M_1 \twoheadrightarrow^\sigma M'_1$.
2. If $M = (\alpha M_1)$, then $M' = (\alpha M'_1)$, where $M_1 \twoheadrightarrow^\sigma M'_1$.
3. If $M = (M_1 \dots M_n)$, then there are standard $\sigma_1, \dots, \sigma_n$ and terms M'_1, \dots, M'_n such that $M_i \twoheadrightarrow^{\sigma_i} M'_i$ ($1 \leq i \leq n$), $M' = (M'_1 \dots M'_n)$ and $\sigma = \sigma_1 \# \dots \# \sigma_n$.
4. If $M = \mu \alpha M_1$ is a $\lambda\mu$ I-term, then $M' = \mu \alpha M'_1$, where $M_1 \twoheadrightarrow^\sigma M'_1$.

Proof By induction on $(|\sigma|, comp(M))$. We consider some typical cases.

1. $M = (\mu \alpha M_1 M_2 \dots M_n)$. Since the head redex $(\mu \alpha M_1 M_2)$ is not involved in σ , point 5 of Definition 3.1 yields the result.
2. $M = \mu \alpha M_1$. If $M = \mu \alpha (\alpha M_2)$ is a θ -redex, then, since M is not involved in σ , point 2 of Definition 3.1 yields that $M_2 \twoheadrightarrow^\sigma N_2$ such that $N = \mu \alpha (\alpha N_2)$. Assume now M is not a θ -redex. Thus M is either not of the form $\mu \alpha (\alpha M_2)$ such that $\alpha \notin Fv(M_2)$ or $M = \mu \alpha (\alpha M_2)$ and $\alpha \in Fv(M_2)$. Then $hd(M) = hd(M_1)$ and, applying the induction hypothesis to M_1 , it is straightforward to check that either M cannot reduce to a term of the form $\mu \alpha (\alpha M'')$, or $M \twoheadrightarrow \mu \alpha (\alpha M'')$ and $\alpha \in Fv(M'')$. Again, by point 2 of Definition 3.1 we obtain the result. \square

We remark that the assumption of M being a $\lambda\mu$ I-term is crucial in Case 4 of Lemma 3.6 as the following example shows.

Example 3.7 *If $M = \mu \alpha (\alpha (\lambda x x (\lambda y x (\alpha x))))$, then $hd(M) = (\lambda x x (\lambda y x (\alpha x)))$. Consider the standard reduction sequence σ : $M = \mu \alpha (\alpha (\lambda x x (\lambda y x (\alpha x)))) \rightarrow_\beta \mu \alpha (\alpha (\lambda x x x)) \rightarrow_\theta (\lambda x x x) = M'$. Then $hd(M)$ is not involved in σ , on the other hand, M' is not of the form $\mu \alpha M'_1$.*

In the next two lemmas our common assumption is that M is a $\lambda\mu$ I-term. The lemmas will serve as auxiliary statements when we prove the uniqueness of the standard normalizing reduction sequence in the case of $\lambda\mu$ I-terms.

Lemma 3.8 *Let M be a $\lambda\mu$ I-term. If $M \twoheadrightarrow^\sigma M'$ is standard such that the head-redex $hd(M)$ of M exists and is not involved in σ , then the head-redex $hd(M')$ of M' exists and it is the unique residual of $hd(M)$ with respect to σ .*

Proof By induction on $(|\sigma|, cxt_y(M))$, taking into account the various cases of Definition 2.9. Let $\sigma = [R] \# \sigma'$. We assume $|\sigma'| = 0$. We examine some of the cases.

1. $M = \mu \alpha M_1$.

- $M = \mu \alpha (\alpha M_2)$.

Assume M is a θ -redex. By assumption, $M_2 \twoheadrightarrow^R M'_2$ such that $M' = \mu \alpha (\alpha M'_2)$. Then our assertion follows. We have also made use of point 2 (b) of Definition 3.1.

Assume now M is not a θ -redex, which implies $\alpha \in Fv(M_2)$. Then $hd(M_1) = hd(M)$ is not R , we have $M_2 \twoheadrightarrow^R M'_2$. Since M is a $\lambda\mu$ I-term, $\alpha \in Fv(M'_2)$ holds. Thus, if $M'_1 = (\alpha M'_2)$, $hd(M'_1) = hd(M')$, by which, and the induction hypothesis, we have the result.

- $M \neq \mu\alpha(\alpha M_2)$. Let $M_1 \rightarrow^R M'_1$, where $M' = \mu\alpha M'_1$. By the induction hypothesis, $hd(M'_1)$ exists and it is the unique residual of $hd(M_1)$, which is $hd(M)$. We prove that $hd(M') = hd(M'_1)$, by which our assertion follows. By Definition 2.9, it is enough to verify that M' is not a θ -redex. Lemma 3.6 shows that the only possibility is $M = \mu\alpha(\alpha M_2)$ for some M_2 provided M' is a θ -redex, but this was excluded by assumption.

2. $M = (\alpha M_1)$.

- $M = (\alpha \mu\beta M_2)$. Then our assumption and point 3 of Definition 3.1 yields the statement.
- M is not a ρ -redex. Then Lemma 3.6 ensures that M' is not a ρ -redex either. If $M_1 \rightarrow^R M'_1$, where $M' = (\alpha M'_1)$, then $hd(M) = hd(M_1)$ and $hd(M') = hd(M'_1)$, by which, together with the induction hypothesis, our claim follows.

3. $M = (\mu\alpha M_1 M_2 \dots M_n)$. Since $hd(M)$ is not R , the only possibility is $M_i \rightarrow^R M'_i$ for some $1 \leq i \leq n$. Hence our assertion follows.

The case $|\sigma'| > 0$ follows by the induction hypothesis. \square

The assumption that M is a $\lambda\mu I$ -term is necessary in the above lemma, too.

Example 3.9 Let $M = \mu\alpha(\alpha (\lambda xx (\lambda yx (\alpha x))))$, then $hd(M) = (\lambda xx (\lambda yx (\alpha x)))$. Consider the reduction sequence $M = \mu\alpha(\alpha (\lambda xx (\lambda yx (\alpha x)))) \rightarrow_\beta \mu\alpha(\alpha (\lambda xx x)) = M'$. In this case $hd(M') = M'$ and it is not a residual of $hd(M)$.

Lemma 3.10 Let M be a $\lambda\mu I$ -term. If $M \twoheadrightarrow^\sigma M'$ is standard and the head-redex $hd(M)$ of M is involved in σ , then $\sigma = [hd(M)]\#\sigma'$ for some σ' .

Proof The proof goes by induction on $|\sigma|$, considering the cases of Definition 3.1. If $|\sigma| = 1$, then the statement is trivial. Assume $\sigma = [R]\#\sigma'$, where $|\sigma'| > 0$, and let $hd(M)$, the head redex of M , be different from R . Let $M \rightarrow^R M'' \twoheadrightarrow^{\sigma'} M'$. By Lemma 3.8, the head redex $hd(M'')$ of M'' exists and it is the unique residual of $hd(M)$ with respect to R . Then $hd(M'')$ is involved in σ' , thus, by the induction hypothesis, we have $\sigma' = [hd(M'')]\#\sigma''$. Now, by examining the various forms of M according to Definition 2.9, we can check easily that the above situation is impossible. \square

Again, $M \in \lambda\mu I$ is necessary for the statement of the previous lemma to hold.

Example 3.11 Let $M = \mu\alpha(\alpha (\lambda xx (\lambda yx (\alpha x))))$, as in Example 3.9, then $hd(M) = (\lambda xx (\lambda yx (\alpha x)))$. Consider the standard reduction sequence σ : $M = \mu\alpha(\alpha (\lambda xx (\lambda yx (\alpha x)))) \rightarrow_\beta \mu\alpha(\alpha (\lambda xx x)) \rightarrow_\theta (\lambda xx x) \rightarrow_\beta x = M'$. Then $hd(M)$ is involved in σ and, on the other hand, σ is not of the form $[hd(M)]\#\sigma'$ for some σ' .

3.2 Calculating the bounds for substitutions

In the following lemmas we examine how standardization is related to substitutions both for λ - and μ -variables. In addition, we give estimations for the lengths of the standard reduction sequences of terms given in the form of substitutions. The lemmas in this subsection are indispensable for proving Lemma 3.20, which is the standardization lemma.

Lemma 3.12 Let $M \twoheadrightarrow^\sigma M'$ be standard, then there exists a $\nu \in St$ such that $M[x := N] \twoheadrightarrow^\nu M'[x := N]$ and $|\nu| = |\sigma|$.

Proof The proof goes by a straightforward induction on $(|\sigma|, \text{comp}(M))$ distinguishing the cases of Definition 3.1. We deal only with the case $M = (\alpha M_1)$. We prove that the choice $\nu = \sigma[x := N]$ is appropriate.

1. If $M \twoheadrightarrow^\sigma M'$ with $M' = (\alpha M'_1)$ and $M_1 \twoheadrightarrow^\sigma M'_1$, then the induction hypothesis applies.
2. (a) If $M \twoheadrightarrow^{\sigma_1} (\alpha \mu\beta P) \rightarrow_\rho P[\beta := \alpha] \twoheadrightarrow^{\sigma_2} M'$ such that $(\alpha \mu\beta P)$ is the first ρ -redex in the sequence, then the induction hypothesis implies that $M[x := N] \twoheadrightarrow^{\sigma_1[x := N]} (\alpha \mu\beta P[x := N])$ and $P[\beta := \alpha][x := N] \twoheadrightarrow^{\sigma_2[x := N]} M'[x := N]$ are standard, moreover, $(\alpha \mu\beta P[x := N])$ is the first ρ -redex in the sequence. We obtain the result immediately from Definition 3.1.
- (b) If $M \twoheadrightarrow^{\sigma_1} (\alpha \mu\beta P) \twoheadrightarrow^{\sigma_2} (\alpha \mu\beta Q) = M'$, where $(\alpha \mu\beta P)$ is the first ρ -redex in the sequence then, by the induction hypothesis, $M[x := N] \twoheadrightarrow^{\sigma_1[x := N]} (\alpha \mu\beta P[x := N])$ and $(\alpha \mu\beta P[x := N]) \twoheadrightarrow^{\sigma_2} (\alpha \mu\beta Q[x := N]) = M'[x := N]$ are standard, which, considering Definition 3.1, yields the result. \square

In the sequel, we make preparations for the estimation of the upper bound of a standard reduction sequence. To this aim, we introduce quantitative notions in relation to reduction sequences.

Definition 3.13 1. Let M be a term and x (resp. α) be a λ -variable (resp. μ -variable). Denote by $|M|_x$ (resp. $|M|_\alpha$) the number of occurrences of x (resp. α) in M .

2. Let σ be the reduction sequence $M \rightarrow^{R_1} M_1 \rightarrow^{R_2} \dots \rightarrow^{R_n} M_n$ and $\alpha \in Fv(M)$. Let $\langle \sigma \rangle_{(\rho, \alpha)}$ denote the number of ρ -reductions of the form $(\alpha \mu\beta P)$ in σ . Furthermore, let $\langle \sigma \rangle_\rho$ be the number of ρ -redexes in σ .

3. If $M \twoheadrightarrow^\sigma M'$, let us denote by $\langle \sigma \rangle_\theta$ the number of θ -redexes in σ .

4. If $x \in Fv(M)$, let us denote by $\text{sumarg}(M, x)$ the sum of the number of arguments of each occurrence of x in M . It is easy to see that $\text{sumarg}(M, x) \leq \text{comp}(M) - 1$.

Lemma 3.14 If $M \twoheadrightarrow^\sigma M'$ is standard and N_1, \dots, N_k are terms for which $\alpha \notin Fv(N_i)$ ($1 \leq i \leq k$), then there exists a standard reduction ν such that $M[\alpha :=_r N_1] \dots [\alpha :=_r N_k] \twoheadrightarrow^\nu M'[\alpha :=_r N_1] \dots [\alpha :=_r N_k]$ and $|\nu| = |\sigma| + k \cdot \langle \sigma \rangle_{(\rho, \alpha)}$.

Proof By induction on $(|\sigma|, \text{comp}(M))$. The case $|\sigma| = 0$ is trivial. If $\sigma = [R] \# \sigma'$, where $M \rightarrow^R M'' \twoheadrightarrow^{\sigma'} M'$, the only interesting case is $M = (\gamma \mu\beta M_1) \rightarrow_\rho M_1[\beta := \gamma] = M'' \twoheadrightarrow^{\sigma'} M'$. If $\alpha \neq \gamma$, then the result follows from the induction hypothesis. Otherwise, we have this reduction (1) : $M[\alpha :=_r N_1] \dots [\alpha :=_r N_k] = (\alpha (\mu\beta M_1[\alpha :=_r N_1] \dots [\alpha :=_r N_k] N_1 \dots N_k)) \twoheadrightarrow_\mu^k (\alpha \mu\beta M_1[\alpha :=_r N_1] \dots [\alpha :=_r N_k][\beta :=_r N_1] \dots [\beta :=_r N_k]) \rightarrow_\rho M_1[\beta := \alpha][\alpha :=_r N_1] \dots [\alpha :=_r N_k]$, from which the estimation for the length of ν follows. Assume σ is standard, we prove by induction on $(|\sigma|, \text{ctxy}(M))$ that ν is standard. We examine the cases of Definition 3.1. We consider only the case when σ is standard by reason of point 3 of Definition 3.1. Let $M = (\gamma M_1)$. If $M_1 \twoheadrightarrow^\sigma M'_1$, the induction hypothesis applies. Otherwise, there are standard σ', σ'' such that $M \twoheadrightarrow^{\sigma'} (\gamma \mu\beta M'') \twoheadrightarrow^{\sigma''} M'$ and $(\gamma \mu\beta M'')$ is the first term in the sequence which is a ρ -redex and either $\sigma'' = [\gamma \mu\beta M''] \# \sigma'''$ for some $\sigma''' \in St$ or $M' = (\gamma \mu\beta M''')$ and $M'' \twoheadrightarrow^{\sigma''} M'''$. Assume $\gamma = \alpha$. Then, by the induction hypothesis, $M[\alpha :=_r N_1] \dots [\alpha :=_r N_k] \twoheadrightarrow^{\nu'} (\alpha (\mu\beta M''[\alpha :=_r N_1] \dots [\alpha :=_r N_k])$

$N_k] N_1 \dots N_k))$ is standard and $(\mu\beta M''[\alpha :=_r N_1] \dots [\alpha :=_r N_k] N_1 \dots N_k)$ is the first μ -redex in the sequence. Then we apply the head reduction sequence of (1), hence Lemma 3.4 involves that ν is standard. If $\gamma \neq \alpha$, then $M[\alpha :=_r N_1] \dots [\alpha :=_r N_k] = (\gamma M_1[\alpha :=_r N_1] \dots [\alpha :=_r N_k]) \rightarrow (\gamma \mu\beta M''[\alpha :=_r N_1] \dots [\alpha :=_r N_k])$. If $\sigma'' = [\gamma \mu\beta M'']\#\sigma'''$, then Lemma 3.5 applies. Otherwise we obtain the result by the induction hypothesis. \square

Lemma 3.15 1. If $N \rightarrow^\sigma N'$ is standard, then there exists a standard reduction ν such that $M[x := N] \rightarrow^\nu M[x := N']$ and $|\nu| \leq |\sigma| \cdot |M|_x + \text{sumarg}(M, x) \cdot (\langle \sigma \rangle_\theta + \langle \sigma \rangle_\rho)$.

2. If $N \rightarrow^\sigma N'$ is standard, then there exists a standard reduction ν such that $M[\alpha :=_r N] \rightarrow^\nu M[\alpha :=_r N']$ and $|\nu| = |\sigma| \cdot |M|_\alpha$.

Proof Let us only deal with Case 1. The proof goes by a straightforward induction on $\text{comp}(M)$. We lean on the points of Definition 2.9. For example, let us consider two of the cases.

1. $M = (x M_1 \dots M_n)$. Let τ_i be the reduction sequences obtained for $M_i[x := N]$ by the induction hypothesis. Let $\tau = \tau_1 \# \dots \# \tau_n$. By induction on $|\sigma|$, we define the following transformation σ° . We eliminate the outermost θ -redexes from σ , that is, redexes R , where $N \rightarrow^{\sigma'} R \rightarrow_\theta R' \rightarrow^{\sigma''} N'$. Observe that an outermost θ -redex appears in σ iff σ is standard by reason of point 2. (a) of Definition 3.1. Let σ be such that $N = \mu\alpha P \rightarrow^{\sigma_1} \mu\alpha(\alpha R) \rightarrow_\theta R \rightarrow^{\sigma_2} N'$, where $\sigma = \sigma_1 \# \sigma_2$ and R is the first θ -redex in σ_1 . Let ξ be $(\mu\alpha P M_1[x := N] \dots M_n[x := N]) \rightarrow_\mu^n \mu\alpha P[\alpha :=_r M_1[x := N]] \dots [\alpha :=_r M_n[x := N]] \rightarrow^{\sigma'_1} \mu\alpha(\alpha (R M_1[x := N] \dots M_n[x := N])) \rightarrow_\theta (R M_1[x := N] \dots M_n[x := N])$, where σ'_1 is obtained from σ_1 by Lemma 3.14. Then let $\sigma^\circ = \xi \# (\sigma_2)^\circ$ where $\sigma = [R] \# \sigma'$. The reduction sequence $M[x := N] \rightarrow^{\sigma^\circ} (N' M_1[x := N] \dots M_n[x := N]) \rightarrow^\tau (N' M_1[x := N'] \dots M_n[x := N'])$ is appropriate. We prove by induction on $|\sigma|$ that $|\sigma^\circ| \leq n + n \cdot (\langle \sigma \rangle_\rho + \langle \sigma \rangle_\theta)$: $|\sigma^\circ| = |\xi| + |\sigma_2^\circ| = 1 + n + |\sigma_1| + n \cdot \langle \sigma_1 \rangle_{(\rho, \alpha)} + |\sigma_2^\circ| \leq 1 + n + |\sigma_1| + n \cdot \langle \sigma_1 \rangle_\rho + |\sigma_2^\circ| \leq 1 + n + |\sigma_1| + n \cdot \langle \sigma_1 \rangle_\rho + |\sigma_2| + n \cdot (\langle \sigma_2 \rangle_\theta + \langle \sigma_2 \rangle_\rho) = |\sigma| + n \cdot (\langle \sigma \rangle_\theta + \langle \sigma \rangle_\rho)$. Then $|\sigma^\circ + \tau| = |\sigma^\circ| + \sum_{i=1}^n |\tau_i| \leq |\sigma| + n \cdot (\langle \sigma \rangle_\theta + \langle \sigma \rangle_\rho) + |\tau_i| \cdot \sum_{i=1}^n |M_i|_x + \sum_{i=1}^n \text{sumarg}(M_i, x) \cdot (\langle \sigma \rangle_\theta + \langle \sigma \rangle_\rho) = |\sigma| \cdot |M|_x + \text{sumarg}(M, x) \cdot (\langle \sigma \rangle_\theta + \langle \sigma \rangle_\rho)$.
2. $M = (\lambda y M_1 M_2 \vec{P})$. The induction hypothesis gives τ_i such that $M_i[x := N] \rightarrow^{\tau_i} M_i[x := N']$ ($1 \leq i \leq k$). Then we can choose $\tau = \tau_1 \# \dots \# \tau_k$. \square

Lemma 3.16 Let $M \rightarrow^\sigma M'$ and $N \rightarrow^\nu N'$ be standard. Then there is a standard reduction τ such that $M[x := N] \rightarrow^\tau M'[x := N']$ and $|\tau| \leq |\sigma| + |M'|_x \cdot |\nu| + \text{sumarg}(M', x) \cdot (\langle \nu \rangle_\theta + \langle \nu \rangle_\rho)$.

Proof The proof goes by induction on $(|\sigma|, \text{comp}(M))$, taking into account the various points of Definition 3.1. The case $|\sigma| = 0$ is treated by Lemma 3.15. Let $\sigma = [R] \# \sigma'$. We treat some typical cases.

1. $M = (\mu\alpha M_1)$. If $M_1 \rightarrow^\sigma M'_1$, then the induction hypothesis applies. Let σ be standard by reason of point 2 (a) of Definition 3.1. Let $M \rightarrow^{\sigma_1} \mu\alpha(\alpha M_2) \rightarrow_\theta M_2 \rightarrow^{\sigma_2} M'$. Lemma 3.12 and the induction hypothesis give standard ν_1 and ν_2 such that $M[x := N] \rightarrow^{\nu_1} \mu\alpha(\alpha M_2[x := N]) \rightarrow_\theta M_2[x := N] \rightarrow^{\nu_2} M'[x := N']$. Moreover, since $\mu\alpha(\alpha M_2[x := N])$ is the first θ -redex in the sequence, $\nu = \nu_1 \# [\mu\alpha(\alpha M_2[x := N])] \# \nu_2$ is standard by virtue of Definition 3.1. The case of point 2. (b) of Definition 3.1 follows from the induction hypothesis.

2. $M = (\mu\alpha M_1 M_2 \dots M_n)$. Assume $M \rightarrow_\mu (\mu\alpha M_1[\alpha :=_r M_2] \dots M_n) \rightarrow^{\sigma'} M'$. Then $(\mu\alpha M_1 M_2)[x := N]$ is the head redex of $M[x := N]$ and Lemma 3.4 together with the induction hypothesis yield the result. If $(\mu\alpha M_1 M_2)$ is not involved in σ , then the induction hypothesis applies.
3. $M = (x M_2 \dots M_n)$. The proof is analogous to that of Lemma 3.15. We define, by induction on ν , a standard reduction sequence ν° in the same way as in Lemma 3.15. We let $\tau = \nu^\circ \# \tau_2 \# \dots \# \tau_n$, where τ_i is obtained from $M_i[x := N]$ ($2 \leq i \leq n$) by the induction hypothesis. By examining the various cases of Definition 3.1, we prove by induction on $|\nu|$ that $\tau \in St$. As to the length of τ , we have $|\tau| = |\nu^\circ| + |\tau_2| + \dots + |\tau_n| \leq |\nu| + (n-1) \cdot (\langle \nu \rangle_\theta + \langle \nu \rangle_\rho) + \sum_{i=2}^n (|\sigma_i| + |M'_i|_x \cdot |\nu| + \text{sumarg}(M'_i, x) \cdot (\langle \nu \rangle_\theta + \langle \nu \rangle_\rho)) = |\sigma| + |M'|_x \cdot |\nu| + \text{sumarg}(M', x) \cdot (\langle \nu \rangle_\theta + \langle \nu \rangle_\rho)$.

The remaining cases are proved analogously. \square

Lemma 3.17 *Let $M \rightarrow^\sigma M'$ and $N \rightarrow^\nu N'$ be standard. Then there is standard sequence τ such that $M[\alpha :=_r N] \rightarrow^\tau M'[\alpha :=_r N']$ and $|\tau| = |\sigma| + \langle \sigma \rangle_{(\rho, \alpha)} + |M'|_\alpha \cdot |\nu|$.*

Proof The proof goes by induction on $(|\sigma|, \text{comp}(M))$, similarly to that of the previous lemma. We consider some of the cases according to Definition 3.1. The case $\sigma=0$ is treated in Lemma 3.15. Let $\sigma = [R] \# \sigma'$ for some σ' .

1. $M = (\beta M_1)$. Assume $\beta = \alpha$. If $M_1 \rightarrow^\sigma M'_1$ with $M' = (\alpha M'_1)$, then the induction hypothesis applies. Otherwise, let $M \rightarrow^{\sigma_1} (\alpha \mu\gamma M_2) \rightarrow_\rho M_2[\gamma := \alpha] \rightarrow^{\sigma_2} M'$. Similarly to the proof of Lemma 3.14, we have the standard reduction sequence $\tau: M[\alpha :=_r N] \rightarrow^{\sigma_1[\alpha :=_r N]} (\alpha (\mu\gamma M_2[\alpha :=_r N] N)) \rightarrow_\mu (\alpha \mu\gamma M_2[\alpha :=_r N][\gamma :=_r N]) \rightarrow_\rho M_2[\gamma := \alpha][\alpha :=_r N] \rightarrow^{\tau_2} M'[\alpha :=_r N']$, where τ_2 is obtained from σ_2 by the induction hypothesis and the standardness follows from Lemma 3.5 and the induction hypothesis, where we have made use of the fact that $(\alpha \mu\gamma M_2)$ is the first ρ -redex in σ_1 . For the length of τ we have $|\tau| = 2 + |\sigma_1| + |\tau_2| = 2 + |\sigma_1| + |\sigma_2| + \langle \sigma_2 \rangle_{(\rho, \alpha)} + |M'|_\alpha \cdot |\nu| = |\sigma| + \langle \sigma \rangle_{(\rho, \alpha)} + |M'|_\alpha \cdot |\nu|$. Assume now $\beta \neq \alpha$. The case when M does not reduce to a ρ -redex or it reduces to a ρ -redex but this is not involved in σ is again obvious. Let $M \rightarrow^{\sigma_1} (\beta \mu\gamma M_2)$. Then $M[\alpha :=_r N] \rightarrow^{\sigma_1[\alpha :=_r N]} (\beta \mu\gamma M_2[\alpha :=_r N]) \rightarrow_\rho M_2[\gamma := \beta][\alpha :=_r N] \rightarrow^{\tau'} M'[\alpha :=_r N']$ is standard, and τ' is obtained by the induction hypothesis. The equation for the length of τ is obviously valid in this case, too.
2. $M = (\mu\beta M_1 M_2 \dots M_n)$. If $(\mu\beta M_1 M_2)$ is not involved in σ , then the induction hypothesis applies. Otherwise, since $hd(M) = hd(M[\alpha :=_r N])$, we have the result by Lemma 3.4 and, again, by the induction hypothesis.

All the remaining cases are proved in a similar way. \square

3.3 The standardization theorem for the $\lambda\mu\rho\theta$ -calculus

We are in a position now to state and prove the standardization theorem for the $\lambda\mu\rho\theta$ -calculus. As an additional result, we obtain an upper bound for the lengths of standard $\lambda\mu$ I-reduction sequences. First of all, we harvest the results of the previous subsection in one definition: the definition below assigns values to pairs formed by redexes and their containing terms. The results of the previous subsection ensure that they will be appropriate for forming an upper bound of the lengths of the standard reduction sequences.

Definition 3.18 Let R be a redex in a term M , the number $m(R, M)$ is defined by

$$m(R, M) = \begin{cases} |M|_\alpha & \text{if } R = (\mu\alpha P Q), \\ 2\text{comp}(M) - 2 & \text{if } R = (\lambda x P Q), \\ 1 & \text{if } R = (\alpha \mu\beta P), \\ 2n - 1 & \text{if } R = \mu\alpha(\alpha P) \text{ and } R \\ & \text{has } n \text{ arguments in } M. \end{cases}$$

The definition for $m(R, M)$ resembles the corresponding definition of Xi in [27], where $m(R)$ is the number of the occurrences of x in P provided $R = (\lambda x P Q)$. The additional redexes, however, compel us to change the value of $m(R, M)$ even for the case of the β -redex. The lemma below will be used in the next subsection.

Lemma 3.19 If R is a redex in M , then $m(R, M) \leq 2(\text{comp}(M) - 1)$.

Proof Immediate by Definition 3.18. \square

The following lemma is the main lemma for obtaining the standardization result and the bound for the standard reduction sequences in Theorem 3.22. In what follows, let $|\sigma|^* = \max(|\sigma|, 1)$, if σ is a reduction sequence.

Lemma 3.20 Let $M \twoheadrightarrow^\sigma M' \rightarrow^R M''$ such that σ is a standard reduction sequence. Then there exists a standard reduction $M \twoheadrightarrow^\tau M''$ such that $|\tau| \leq 1 + \max(m(R, M'), 2) \cdot |\sigma|^*$. Furthermore, if M is a $\lambda\mu$ I-term, then $1 + |\sigma| \leq |\tau|$.

Proof The proof goes by induction $(|\sigma|, \text{comp}(M))$. The case of $|\sigma| = 0$ is obvious, thus we may assume $|\sigma| > 0$. We examine the points of Definition 3.1. We treat some of the more interesting cases.

1. $M = (\alpha M_1)$. If $M_1 \twoheadrightarrow^\sigma M'_1$ such that $M' = (\alpha M'_1)$ and there are no ρ -redexes as head redexes in σ including M'' , then the induction hypothesis applies. Assume $M \twoheadrightarrow^{\sigma_1} (\alpha \mu\beta M_2)$ such that $\sigma_1 \in St$ and $(\alpha \mu\beta M_2)$ is the first ρ -redex in the sequence. Let us suppose, according to point 3 (a) of Definition 3.1, $(\alpha \mu\beta M_2) \rightarrow_\rho M_2[\beta := \alpha] \twoheadrightarrow^{\sigma_2} M'$, where $\sigma = \sigma_1 \# [(\alpha \mu\beta M_2)] \# \sigma_2$ and $\sigma_i \in St$ ($i \in \{1, 2\}$). By the induction hypothesis applied to σ_2 , we obtain a $\tau' \in St$ such that $|\tau'| \leq 1 + \max(m(R, M'), 2) \cdot |\sigma_2|^*$. Then let $\tau = \sigma_1 \# [(\alpha \mu\beta M_2)] \# \tau'$. Hence $|\tau| = 1 + |\sigma_1| + |\tau'| \leq 1 + |\sigma_1| + 1 + \max(m(R, M'), 2) \cdot |\sigma_2|^* \leq 1 + \max(m(R, M'), 2) \cdot |\sigma|^*$. Assume we have $(\alpha \mu\beta M_2) \twoheadrightarrow^{\sigma_2} (\alpha \mu\beta M'_2)$ with $M' = (\alpha \mu\beta M'_2)$ and $M_2 \twoheadrightarrow^{\sigma_2} M'_2$, by reason of point 3 (b) of Definition 3.1. If $R \leq M'_2$, then we obtain the result by the induction hypothesis. Assume $R = M'$. Then $\tau = \sigma_1 \# [(\alpha \mu\beta M_2)] \# \sigma'_2$ is appropriate, where σ'_2 is $M_2[\beta := \alpha] \twoheadrightarrow^{\sigma_2[\beta := \alpha]} M'_2[\beta := \alpha]$. The estimation for $|\tau|$ follows easily, since $|\sigma'_2| = |\sigma_2|$. Finally, if M is a $\lambda\mu$ I-term, the result follows from the induction hypothesis by inspection of the various subcases. For example, consider the case when $(\alpha \mu\beta M_2) \rightarrow_\rho M_2[\beta := \alpha] \twoheadrightarrow^{\sigma_2} M'$, where $\sigma = \sigma_1 \# [(\alpha \mu\beta M_2)] \# \sigma_2$, that is, the case described by point 3 (a) of Definition 3.1. If τ' is the standard reduction sequence corresponding to σ_2 by the induction hypothesis and $\tau = \sigma_1 \# [(\alpha \mu\beta M_2)] \# \tau'$, then $1 + |\sigma_2| \leq |\tau'|$ and we obtain the result.
2. $M = (\mu\alpha M_1 M_2 \dots M_n)$. Let σ be standard by virtue of point 5.(a) of Definition 3.1. Then $(\mu\alpha M_1 M_2 \dots M_n) \rightarrow_\mu (\mu\alpha M_1[\alpha :=_r M_2] \dots M_n) \twoheadrightarrow^{\sigma'} M'$ with $\sigma' \in St$. The induction hypothesis applied to σ' provides us with a standard τ' with appropriate length such that $(\mu\alpha M_1[\alpha :=_r M_2] \dots M_n) \twoheadrightarrow^{\tau'} M''$. By this the result follows. Assume σ is standard by reason of point 5.(b) of Definition 3.1. Then $(\mu\alpha M_1 M_2 \dots M_n) \rightarrow^{\sigma_1} (\mu\alpha M'_1 M_2 \dots M_n) \twoheadrightarrow^{\sigma_2} \dots \twoheadrightarrow^{\sigma_n} (\mu\alpha M'_1 M'_2 \dots M'_n) = M'$, where $\sigma = \sigma_1 \# \dots \# \sigma_n$. If $R \leq M'_i$,

then the induction hypothesis gives the result. Let $R = (\mu\alpha M'_1 M'_2)$. Then τ can be chosen as $(\mu\alpha M_1 M_2 \dots M_n) \rightarrow_\mu (\mu\alpha M_1 [\alpha :=_r M_2] \dots M_n) \rightarrow^{\tau_1} (\mu\alpha M'_1 [\alpha :=_r M'_2] \dots M'_n) \rightarrow^{\sigma_3} \dots \rightarrow^{\sigma_n} (\mu\alpha M'_1 [\alpha :=_r M'_2] \dots M'_n)$, where τ_1 is obtained from Lemma 3.17. Moreover, $|\tau| = 1 + |\tau_1| + \sum_{i=3}^n |\sigma_i| = 1 + |\sigma_1| + \langle \sigma_1 \rangle_{(\rho, \alpha)} + |M'|_\alpha \cdot |\sigma_2| + \sum_{i=3}^n |\sigma_i| \leq 1 + 2|\sigma_1| + |M'|_\alpha \cdot |\sigma_2| + \sum_{i=3}^n |\sigma_i| \leq 1 + \max(m(R, M'), 2) \cdot |\sigma|^*$. Assume $R = \mu\alpha(\alpha M''_1)$ is a θ -redex. In this case σ_1 is standard by virtue of point 2 (b) (ii) of Definition 3.1. Let $\mu\alpha(\alpha M''_1)$ be the first θ -redex such that an initial segment σ'_1 of σ_1 produces $\mu\alpha(\alpha M''_1)$ starting from $\mu\alpha M_1$. Let $\sigma_1 = \sigma'_1 \# \sigma''_1$. Then $(\mu\alpha M_1 M_2 \dots M_n) \rightarrow_\mu^{n-1} \mu\alpha M_1 [\alpha :=_r M_2] \dots [\alpha :=_r M_n] \rightarrow^{\tau_1} \mu\alpha(\alpha (M''_1 M_2 \dots M_n)) \rightarrow_\theta (M''_1 M_2 \dots M_n) \rightarrow^{\sigma''_1} (M'_1 M_2 \dots M'_n) \rightarrow^{\sigma_2} \dots \rightarrow^{\sigma_n} (M'_1 M'_2 \dots M'_n) = M'$ is standard, where τ_1 is obtained from σ'_1 by Lemma 3.14. As to the length of τ , we have $|\tau| = 1 + (n-1) + |\tau_1| + |\sigma''_1| + \sum_{i=2}^n |\sigma_i| = 1 + (n-1) + |\sigma'_1| + (n-1) \cdot \langle \sigma'_1 \rangle_{(\rho, \alpha)} + |\sigma''_1| + \sum_{i=2}^n |\sigma_i| \leq 1 + |\sigma| + (n-1) \cdot (1 + |\sigma|) = 1 + n \cdot |\sigma| + (n-1) \leq 1 + \max(m(R, M'), 2) \cdot |\sigma|^*$. When M is a $\lambda\mu$ I-term, we obtain the result by the induction hypothesis. Let us only treat the last case, where $(\mu\alpha M_1 M_2 \dots M_n) \rightarrow^{\sigma_1} (\mu\alpha M'_1 M_2 \dots M_n) \rightarrow^{\sigma_2} \dots \rightarrow^{\sigma_n} (\mu\alpha M'_1 M'_2 \dots M'_n) = M'$ and $R = \mu\alpha M'_1 = \mu\alpha(\alpha M''_1)$. If $\mu\alpha(\alpha M''_1)$ is the first θ -redex in σ_1 such that $\sigma_1 = \sigma'_1 \# \sigma''_1$ and τ_1 is obtained from σ'_1 by Lemma 3.14 and τ is defined as above, then $1 + |\sigma| = 1 + |\sigma'_1| + |\sigma''_1| + \sum_{i=2}^n |\sigma_i| \leq |\tau| = (n-1) + |\tau_1| + 1 + |\sigma''_1| + \sum_{i=2}^n |\sigma_i|$, where $|\sigma'_1| \leq |\tau_1|$ by Lemma 3.14. \square

Definition 3.21 Let σ be the reduction sequence $M_1 \rightarrow^{R_1} M_2 \rightarrow^{R_2} \dots \rightarrow^{R_n} M_{n+1}$. Denote by $\mathcal{M}(\sigma)$ (the measure of σ) the number $\prod_{i=1}^n (1 + \max(m(R_i, M_i), 2))$.

Theorem 3.22 Let σ be the reduction sequence $M = M_1 \rightarrow^{R_1} M_2 \rightarrow^{R_2} \dots \rightarrow^{R_n} M_{n+1}$. Then there is a standard reduction sequence $st(\sigma)$ such that $M_1 \rightarrow^{st(\sigma)} M_{n+1}$ and $|st(\sigma)| \leq \mathcal{M}(\sigma)$. Moreover, if M is a $\lambda\mu$ I-term, then $|\sigma| \leq |st(\sigma)|$ also holds.

Proof The statement of the theorem is proved by induction on $|\sigma|$.

1. If $|\sigma| = 1$, then our claim follows directly from Lemma 3.20.
2. Let $\sigma = \sigma' \# [R_n]$, where $|\sigma'| \geq 1$. By the induction hypothesis, we can find a standard $st(\sigma')$ with appropriate length such that $M_1 \rightarrow^{st(\sigma')} M_n$. Moreover, $|st(\sigma')|^* = |st(\sigma')|$. Then, by Lemma 3.20, there is a standard $M_1 \rightarrow^\tau M_{n+1}$ such that $|\tau| \leq 1 + \max(m(R_n, M_n), 2) \cdot |st(\sigma')|^* \leq (1 + \max(m(R_n, M_n), 2)) \cdot |st(\sigma')|^*$, which yields the result. \square

Theorem 3.23 If M is a $\lambda\mu$ I-term, then the standard reduction sequence starting from M and leading to the normal form of M is the leftmost reduction sequence and it is a reduction sequence of maximal length.

Proof Let M be a $\lambda\mu$ I-term. Assume $M \rightarrow^\sigma M'$ where M' is the normal form of M . The proof goes by induction on $(|\sigma|, cxt_y(M))$. We may assume $M \in HNF$. Otherwise, by Lemma 3.8, the head redex of M is involved in σ , then Lemma 3.10 yields that $\sigma = [hd(M)] \# \sigma'$. That is, if $M \notin HNF$, then the induction hypothesis applies. Let $M = (x M_2 \dots M_n)$. By Definition 3.1 there exist $\sigma_i \in St$ ($2 \leq i \leq n$) such that $\sigma = \sigma_2 \# \dots \# \sigma_n$ and $(x M_2 \dots M_n) \rightarrow^{\sigma_2} (x M'_2 \dots M_n) \rightarrow^{\sigma_3} \dots \rightarrow^{\sigma_n} (x M'_2 \dots M'_n)$. Then the induction hypothesis applied to σ_i ($2 \leq i \leq n$) gives the result. The leftmost reduction has a maximal length by Theorem 3.22. \square

4 The estimation for the $\lambda\mu\rho\theta$ -calculus

In this section we present an application of Theorem 3.22, which, together with the existence of a standard reduction sequence, even provides us with a bound for the length of the sequence. Making use of the fact that, by Theorem 3.23, the standard reduction sequence for a $\lambda\mu I$ -term is unique, we find upper bounds for reduction sequences of $\lambda\mu I$ -terms first by choosing a normalization sequence σ the measure of which, $\mathcal{M}(\sigma)$, can be easily estimated from above. By Theorem 3.22, we obtain then that $|\sigma|$ obeys that bound, too. We extend this result to the general case by finding a translation $[M]_k$ of M with an appropriate k , where $[M]_k$ is a $\lambda\mu I$ -term such that lengths of the types of the redexes in M is the same as those of $[M]_k$ and $\eta(M) \leq \eta([M]_k)$ and the complexity of $[M]_k$ is bounded by a linear function of the complexity of M .

4.1 The estimation for the $\lambda\mu\rho\theta I$ -calculus

In this section we give an estimation for the lengths of the reduction sequences in the $\lambda\mu\rho\theta I$ -calculus. To this end we define a normalization strategy such that the lengths of reduction sequences obeying that strategy can be assessed from above and we can even establish bounds for the sizes of the developments. Prior to this, we need the rank of a redex.

Definition 4.1 1. The rank of a redex R in a term M is

$$\text{rank}(R, M) = \begin{cases} lh(\text{type}(\lambda x M_1)) & \text{if } R = (\lambda x M_1 M_2), \\ lh(\text{type}(\mu \alpha M_1)) & \text{if } R = (\mu \alpha M_1 M_2), \\ lh(\text{type}(\mu \beta M)) & \text{if } R = (\alpha \mu \beta M), \\ lh(\text{type}(\mu \alpha (\alpha M))) & \text{if } R = \mu \alpha (\alpha M). \end{cases}$$

2. The rank of a term M is $\text{rank}(M) = \max\{\text{rank}(R, M) \mid R \text{ is a redex in } M\}$.

3. Define $NF_k = \{M \mid \text{rank}(M) < k\}$.

The following lemma states that reductions do not decrease the rank.

Lemma 4.2 Let M, N be terms.

1. We have $\text{rank}(M[x := N]) \leq \max\{\text{rank}(M), \text{rank}(N), lh(\text{type}(x))\}$ and $\text{rank}(M[\alpha :=_r N]) \leq \max\{\text{rank}(M), \text{rank}(N), lh(\text{type}^*(\alpha))\}$, where $\text{type}^*(\alpha) = A$ if $\text{type}(\alpha) = \neg A$.

2. If $M \rightarrow M'$, then $\text{rank}(M) \geq \text{rank}(M')$.

Proof

1. By induction on $\text{comp}(M)$.

2. It is enough to prove if $M \rightarrow^R M'$, then $\text{rank}(M) \geq \text{rank}(M')$. The proof goes by induction on $\text{comp}(M)$ and we use the first item. \square

Definition 4.3 1. We say that a reduction sequence ν is a k -reduction sequence, if every redex in ν is of rank k .

2. A reduction sequence $M \rightarrow^\nu M'$ is a k -normalization for a given term M , if it is a k -reduction sequence and $M' \in NF_k$.
3. A reduction sequence ξ starting from a term is good, if, at each reduction step, it chooses the leftmost, innermost redex of maximal rank, that is, the redex containing no other redexes of maximal rank and stands in the leftmost position among these redexes.

Let σ be a good reduction sequence starting from M , assume $\text{rank}(M) = k$. Then σ starts with the leftmost, innermost redex of rank k and chooses the leftmost, innermost redex of maximal rank every time. Since M is strongly normalizable, σ is necessarily finite. By Lemma 4.2, the ranks of the redexes involved in σ form a monotone decreasing sequence. Thus, if σ is a good normalizing sequence, then the sequence of redexes of rank k in σ comes to an end and σ continues with a leftmost, innermost redex of maximal rank. Hence, σ is the concatenation of l_i -normalization sequences ($1 \leq i \leq s$) with $l_1 = k > l_2 > \dots > l_s \geq 1$.

The next two lemma show that good k -normalization sequences can be dissected easily so that we are able to estimate their lengths in the proof of Lemma 4.6.

Lemma 4.4

1. Let $\text{rank}(\mu\alpha P Q) = k$ and $x \notin Fv(P)$. If $(\mu\alpha P Q) \rightarrow^\nu U$ and ν is a good k -normalization sequence, there are terms P', Q', U' and good k -normalization sequences ν_1, ν_2, ν_3 such that $P \rightarrow^{\nu_1} P', Q \rightarrow^{\nu_2} Q', (\mu\alpha P' x) \rightarrow^{\nu_3} U', U = U'[x = Q']$ and $\nu = \nu_1 \# \nu_2 \# \nu_3[x := Q']$.
2. Let $\text{rank}(\lambda y P Q) = k$ and $x \notin Fv(P)$. If $(\lambda y P Q) \rightarrow^\nu U$ and ν is a good k -normalization sequence, there are terms P', Q', U' and good k -normalization sequences ν_1, ν_2 such that $P \rightarrow^{\nu_1} P', Q \rightarrow^{\nu_2} Q', (\lambda y P' x) \rightarrow^{\nu_3} P'[y := x] = P'', U = P''[x := Q']$ and $\nu = \nu_1 \# \nu_2 \# \nu_3[x := Q']$.

Proof

1. The algorithm proceeds by eliminating the innermost k -redexes from left to right, that is we have (possibly empty) ν_1 and ν_2 - both being k -normalization sequences such that $\nu_1 \# \nu_2$ is an initial subsequent of ν and $P \rightarrow^{\nu_1} P' \in NF_k, Q \rightarrow^{\nu_2} Q' \in NF_k$. Then ν continues with reducing $(\mu\alpha P' Q')$ and the redexes created by this reduction. It is immediate to check that when reducing $(\mu\alpha P' Q')$, the created k -redexes can only be redexes of the form $(\lambda y V[\alpha :=_r Q'] Q')$ for some $\lambda y V$ of rank k such that $(\alpha \lambda y V) \leq P'$, so for every k -redex R in $\mu\alpha P'[\alpha :=_r Q']$ there is an R' in $\mu\alpha P'[\alpha :=_r x]$ such that $R = R'[x := Q']$. Reducing with these β -redexes in $\mu\alpha P'[\alpha :=_r Q']$, no more k -redexes are created. This proves our assertion.
2. Analogous to the first point. □

Lemma 4.5 Let $\text{rank}((\mu\alpha P x)) = k, \mu\alpha P \in NF_k$ and $x \notin Fv(P)$. If $(\mu\alpha P x) \rightarrow^\nu U$, ν is a good k -normalization sequence, and $U \in NF_k$, then $|\nu| \leq \text{comp}(P)$ and $\text{comp}(U) \leq 2 \cdot \text{comp}(P)$.

Proof Since $\mu\alpha P \in NF_k$, in $\mu\alpha P[\alpha :=_r x]$ k -redexes of the form $(\lambda y Q[\alpha :=_r x] x)$ can only occur, where $(\alpha \lambda y Q) \leq P$ and $\text{rank}(\lambda y Q) = k$. Subsequently reducing these redexes gives U , which means that U can be obtained in at most $|P|_\alpha + 1 \leq \text{comp}(P)$ steps. Considering the above argument, since x is a variable, the β -reduction steps in ν does not increase the size of the term, so $\text{comp}(U) \leq \text{comp}(\mu\alpha P[\alpha :=_r x]) = 1 + \text{comp}(P) + |P|_\alpha \leq 2 \cdot \text{comp}(P)$. □

The lemma below gives estimations for good k -normalization sequences. We may observe that the obtained bounds does not depend on k .

Lemma 4.6 *Let M be a term such that $\text{rank}(M) = k$. If $M \rightarrow^\nu M'$ and ν is a good k -normalization sequence, then $\text{comp}(M') \leq 2^{\text{comp}(M)-1}$ and $|\nu| \leq 2^{\text{comp}(M)-1}$.*

Proof The proof of $\text{comp}(M') \leq 2^{\text{comp}(M)-1}$ goes by induction on $\text{comp}(M)$.

1. The case $M = x$ or $M = \lambda x M_1$ is obvious.
2. Let $M = \mu \alpha M_1$.
 - (a) If $M = \mu \alpha(\alpha M_1)$ is a θ -redex of rank k , then, since the algorithm eliminates k -redexes from bottom to up and from left to right, we have a $\nu' \leq \nu$ such that $\mu \alpha(\alpha M_1) \rightarrow^{\nu'} \mu \alpha(\alpha M'_1) \rightarrow^R M'_1 = M'$. But in this case $M \rightarrow_\theta M_1 \rightarrow^{\nu'} M'$ is valid as well, thus by the induction hypothesis $\text{comp}(M') \leq 2^{\text{comp}((M_1)-1)} < 2^{\text{comp}(M)-1}$.
 - (b) If $\mu \alpha M_1$ is not a θ -redex, but reduces to a θ -redex of rank k in the course of the process, then a reasoning analogous to the above one works.
 - (c) If $\mu \alpha M_1$ is not a θ -redex and it neither reduces to a θ -redex, then the induction hypothesis applies.
3. Let $M = (M_1 M_2)$.
 - (a) If M is not a k -redex, then we prove that M cannot reduce to a k -redex. Suppose on the contrary that there is some initial subsequent of ν such that it reduces M to a k -redex, take ν' as the shortest such reduction sequence. Suppose M reduces to a μ -redex (the case of a β -redex is similar). In this case we have $M \rightarrow^{\nu'} (\mu \beta N_1 N_2)$, where $M_1 \rightarrow \mu \beta N_1$ and $M_2 \rightarrow N_2$. Then $M \rightarrow^{\nu''} (N_3 N_2) \rightarrow^{R'} (\mu \beta N_1 N_2)$ must hold for some R' , ν'' such that $\nu' = \nu'' \# [R']$ and for some N_3 , N_3 not beginning with a μ . This means $N_3 = R'$ would be again a k -redex, but a straightforward examination of the possible cases shows it is impossible. Hence we have $M' = (M'_1 M'_2)$, $\nu = \nu_1 \# \nu_2$ for some k -reduction sequences ν_1 , ν_2 and $M_i \rightarrow^{\nu_i} M'_i$ ($i \in \{1, 2\}$). Thus by the induction hypothesis $\text{comp}(M') = \text{comp}(M'_1) + \text{comp}(M'_2) \leq 2^{\text{comp}(M_1)-1} + 2^{\text{comp}(M_2)-1} \leq 2^{\text{comp}(M)-1}$.
 - (b) If M is a k -redex and $M = (\mu \alpha M_1 M_2)$, then M is involved in ν as a μ -redex. By Lemma 4.4, we have M'_1 , M'_2 , M'' and ν_1 , ν_2 , ν_3 such that $M_1 \rightarrow^{\nu_1} M'_1$, $M_2 \rightarrow^{\nu_2} M'_2$, $(\mu \alpha M'_1 x) \rightarrow^{\nu_3} M''$, $M' = M''[x := M'_2]$ and $\nu = \nu_1 \# \nu_2 \# \nu_3[x := M'_2]$, provided $x \notin Fv(M_1)$. From this, by Lemma 4.5 and by the induction hypothesis, $\text{comp}(M') = \text{comp}(M''[x := M'_2]) = \text{comp}(M'') + |M''|_x \cdot (\text{comp}(M'_2) - 1) < \text{comp}(M'') \cdot \text{comp}(M'_2) \leq 2 \cdot \text{comp}(M'_1) \cdot \text{comp}(M'_2) \leq 2 \cdot 2^{\text{comp}(M_1)-1} \cdot 2^{\text{comp}(M_2)-1} < 2^{\text{comp}(M)-1}$. The case $M = (\lambda x M_1 M_2)$ is similar.
4. Let $M = (\alpha M_1)$.
 - (a) If M does not reduce to a k -redex, then the result is obvious.
 - (b) If M is either a k -redex, or reduces to a k -redex, then there is a ν' and a $\mu \beta M_2 \in NF_k$ such that $(\alpha M_1) \rightarrow^{\nu'} (\alpha \mu \beta M_2) \rightarrow^R M_2[\beta := \alpha]$ and $\nu' \# [R] = \nu$. The induction hypothesis for M_1 gives the result.

We prove $|\nu| \leq 2^{\text{comp}(M)-1}$ by induction on $\text{comp}(M)$. The only interesting case is when M is a redex of rank k . Let, for example, $M = (\mu\alpha M_1 M_2)$. Since ν is a k -normalization sequence we can assume again that M is involved in ν . By Lemma 4.4, we have M'_1, M'_2 and k -normalization sequences ν_1, ν_2, ν_3 such that $M_1 \rightarrow^{\nu_1} M'_1$, $M_2 \rightarrow^{\nu_2} M'_2$, $(\mu\alpha M'_1 x) \rightarrow^{\nu_3} M''$, $M' = M''[x = M'_2]$ and $\nu = \nu_1 \# \nu_2 \# \nu_3[x := M'_2]$, provided $x \notin Fv(M_1)$. Then, using Lemma 4.5 and the induction hypothesis, we obtain $|\nu| = |\nu_1| + |\nu_2| + |\nu_3[x := M'_2]| = |\nu_1| + |\nu_2| + |\nu_3| \leq 2^{\text{comp}(M_1)-1} + 2^{\text{comp}(M_2)-1} + 2^{\text{comp}(M_1)-1} = 2^{\text{comp}(M_1)} + 2^{\text{comp}(M_2)-1} \leq 2^{\text{comp}(M)-1}$. \square

Definition 4.7 Let tower defined by $\text{tower}(n, m) = \begin{cases} m & \text{if } n = 0, \\ 2^{\text{tower}(n-1, m)} & \text{if } n > 0. \end{cases}$

Theorem 4.8 Let M be a term such that $\text{rank}(M) = k$. If $M \rightarrow^\sigma N$, σ is a good reduction sequence and $N \in NF$, then $\mathcal{M}(\sigma) < \text{tower}(k+1, \text{comp}(M))$.

Proof We first prove by induction on k that $\mathcal{M}(\sigma) < \text{tower}(1, \text{tower}(1, \text{comp}(M))) + \sum_{i=2}^k \text{tower}(i, \text{comp}(M) - 1)$.

1. If $k = 1$, then σ is a 1-normalization sequence. Suppose σ is $M = M_1 \rightarrow^{R_1} M_2 \rightarrow^{R_2} \dots \rightarrow^{R_{n-1}} M_n \rightarrow^{R_n} M_{n+1}$ for some $n \geq 1$. We have, by Lemmas 3.19 and 4.6, $1 + \max(m(R_i, M_i), 2) \leq 2^{\text{comp}(M_i)} - 1 \leq 2 \cdot 2^{\text{comp}(M)-1} - 1 < 2^{\text{comp}(M)}$, then $\mathcal{M}(\sigma) = \prod_{i=1}^n (1 + \max(m(R_i, M_i), 2)) < \prod_{i=1}^n 2^{\text{comp}(M)} = 2^{n \cdot \text{comp}(M)}$. Again, by Lemma 4.6, we obtain $n = |\sigma| \leq 2^{\text{comp}(M)-1}$, so $\mathcal{M}(\sigma) < 2^{\text{comp}(M) \cdot 2^{\text{comp}(M)-1}} \leq 2^{2^{\text{comp}(M)}} = \text{tower}(1, \text{tower}(1, \text{comp}(M)))$.
2. Let $\text{rank}(M) = k+1$ and $k \geq 1$. Assume $M \rightarrow^{\sigma'} M' \rightarrow^{\sigma''} N \in NF$, where σ' is a $k+1$ -normalization sequence starting from M . By the induction hypothesis, we have $\mathcal{M}(\sigma'') < \text{tower}(1, \text{tower}(1, \text{comp}(M'))) + \sum_{i=2}^k \text{tower}(i, \text{comp}(M') - 1)$. As above, we obtain again $\mathcal{M}(\sigma') < 2^{2^{\text{comp}(M)}}$. Then, using the multiplicity of \mathcal{M} and Lemma 4.6, we can assert $\mathcal{M}(\sigma) = \mathcal{M}(\sigma') \cdot \mathcal{M}(\sigma'') < 2^{2^{\text{comp}(M)}} \cdot \text{tower}\left(1, \text{tower}(1, \text{comp}(M')) + \sum_{i=2}^k \text{tower}(i, \text{comp}(M') - 1)\right) < 2^{2^{\text{comp}(M)}} \cdot \text{tower}\left(1, \text{tower}(1, \text{tower}(1, \text{comp}(M) - 1)) + \sum_{i=2}^k \text{tower}(i, \text{tower}(1, \text{comp}(M) - 1))\right) = 2^{2^{\text{comp}(M)}} \cdot 2^{\underbrace{2^{2^{\text{comp}(M)-1}} + \dots + 2^{2^{\text{comp}(M)-1}}}_k} = \text{tower}(1, \text{tower}(1, \text{comp}(M))) + \sum_{i=2}^{k+1} \text{tower}(i, \text{comp}(M) - 1)$.

Finally, we prove by induction on k that

$$\text{tower}(1, \text{comp}(M)) + \sum_{i=2}^k \text{tower}(i, \text{comp}(M) - 1) \leq \text{tower}(k, \text{comp}(M)).$$

The case $k = 1$ is obvious. Let $k = n+1$ and $n \geq 1$. Applying the induction hypothesis, we obtain $\text{tower}(1, \text{comp}(M)) + \sum_{i=2}^{n+1} \text{tower}(i, \text{comp}(M) - 1) =$

$$\underbrace{2^{\text{comp}(M)} + 2^{2^{\text{comp}(M)-1}} + \dots + 2^{2^{\text{comp}(M)-1}}}_{n+1} \leq$$

$$\text{tower}(n, \text{comp}(M)) + \text{tower}(n+1, \text{comp}(M) - 1) < \text{tower}(n+1, \text{comp}(M)). \quad \square$$

Corollary 4.9 Let M be a $\lambda\mu I$ -term of rank k . Every reduction sequence starting from M has length less than $\text{tower}(k+1, \text{comp}(M))$.

Proof Let N be the normal-form of M . By Definition 4.3 and Theorem 4.8, there exists a σ such that $M \rightarrow^\sigma N$ and $\mathcal{M}(\sigma) < \text{tower}(k+1, \text{comp}(M))$. By Theorem 3.22, there is a standard σ' such that $M \rightarrow^{\sigma'} N$ and $|\sigma'| < \mathcal{M}(\sigma)$. The result follows now from Theorem 3.23. \square

4.2 Some properties of the function η

In the next subsection we undertake the task of estimating the lengths of reduction sequences by transforming the starting terms into $\lambda\mu I$ -terms. In order to make the estimation work, we have to prove that the longest reduction sequences of the transformed terms are at least as long as those of the original terms. To this end, we give some estimations concerning longest reduction sequences of terms and their reducts. This subsection prepares the treatment of the general case in the next subsection.

Lemma 4.10 *Let M , N and \vec{P} be $\lambda\mu I$ -terms. If $\alpha \notin Fv(N)$, then $\eta((\mu\alpha\langle M, (\alpha z) \rangle \vec{P})) + \eta(N) \leq \eta((\mu\alpha\langle M, (\alpha (z N)) \rangle \vec{P}))$.*

Proof Let $U = (\mu\alpha\langle M, (\alpha z) \rangle \vec{P})$, $V = (\mu\alpha\langle M, (\alpha (z N)) \rangle \vec{P})$. If \vec{P} is empty, the result is trivial, so may assume \vec{P} is not empty and its components are M_1, \dots, M_n . We are going to prove if $U \rightarrow^{\sigma_1} U'$, $N \rightarrow^{\sigma_2} N'$ for some σ_1 , σ_2 , U' , N' , then we have a reduction sequence ν of V such that $|\sigma_1| + |\sigma_2| \leq |\nu|$. By the second part of Theorem 3.22, it is enough to restrict our attention to the case when σ_1 and σ_2 are standard. We may assume that the head-redex of U is involved in σ_1 , otherwise the result is trivial. Furthermore, we may suppose that $\mu\alpha\langle M, (\alpha z) \rangle$ is reduced in $|\sigma_1|$ with all of its arguments M_1, \dots, M_n . Then σ_1 is of the form $U \rightarrow^\xi \mu\alpha\langle M[\alpha :=_r M_1] \dots [\alpha :=_r M_n], (\alpha (z M_1 \dots M_n)) \rangle \rightarrow^\zeta \mu\alpha\langle M', (\alpha (z M_1 \dots M_n)) \rangle \rightarrow^{\zeta^*} \mu\alpha\langle M', (\alpha (z M'_1 \dots M'_n)) \rangle$, where $M[\alpha :=_r M_1] \dots [\alpha :=_r M_n] \rightarrow^\zeta M'$ and $\zeta^* = \zeta_1 \# \dots \# \zeta_n$ with $M_i \rightarrow^{\zeta_i} M'_i$ for $1 \leq i \leq n$. Let ξ' be $V \rightarrow^{\xi'} \mu\alpha\langle M[\alpha :=_r M_1] \dots [\alpha :=_r M_n], (\alpha (z N M_1 \dots M_n)) \rangle$, then choosing ν as $\nu = \xi' \# \zeta \# \sigma_2 \# \zeta^*$ is appropriate. \square

Lemma 4.11 *Let $M = (\lambda x M_1 M_2 \vec{P})$ and $N = (M_1[x := M_2] \vec{P}')$.*

1. *If $x \in Fv(M_1)$ and N is strongly normalizable, then M is also strongly normalizable and $\eta(M) = \eta(N) + 1$.*
2. *If $x \notin Fv(M_1)$ and N, M_2 are strongly normalizable, then M is also strongly normalizable and $\eta(M) = \eta(N) + \eta(M_2) + 1$.*

Proof

1. Let $M \rightarrow^\sigma U$ be an arbitrary reduction sequence, we are going to show that $|\sigma| \leq \eta(N) + 1$, from which the result follows. We may suppose that $(\lambda x M_1 M_2)$ is involved in σ . Then σ is of the following form for some σ_1 and σ_2 , $M = (\lambda x M_1 M_2 \vec{P}) \rightarrow^{\sigma_1} M' = (\lambda x M'_1 M'_2 \vec{P}') \rightarrow (M'_1[x := M'_2] \vec{P}') \rightarrow^{\sigma_2} U$ where $M_i \rightarrow^{\nu_i} M'_i$ ($i \in \{1, 2\}$), $\vec{P} \rightarrow^{\nu_3} \vec{P}'$ and $\sigma_1 = \nu_1 \# \nu_2 \# \nu_3$. Let σ' denote the reduction sequence $M = (\lambda x M_1 M_2 \vec{P}) \rightarrow N = (M_1[x := M_2] \vec{P}) \rightarrow^{\sigma^*} U$, where $\sigma^* = \nu'_1 \# \nu'_2 \# \nu_3 \# \sigma_2$ and ν'_1 is constructed from ν_1 by Lemma 3.12 with $M_1 \rightarrow^{\nu_1} M'_1$ and M_2 and ν'_2 is obtained by applying Lemma 3.15 to M_1 and $M_2 \rightarrow^{\nu_2} M'_2$. Then $|\sigma| \leq \eta(N) + 1$, which is the desired result.
2. Let $M \rightarrow^\sigma U$ be an arbitrary reduction sequence, it is enough to show that $|\sigma| \leq \eta(N) + \eta(M_2) + 1$. We may suppose that $(\lambda x M_1 M_2)$ is involved in σ . Then σ is of the form $M = (\lambda x M_1 M_2 \vec{P}) \rightarrow^{\sigma_1} M' = (\lambda x M'_1 M'_2 \vec{P}') \rightarrow_\beta (M'_1 \vec{P}') \rightarrow^{\sigma_2} U$ where $M_i \rightarrow^{\nu_i} M'_i$ ($i \in \{1, 2\}$), $\vec{P} \rightarrow^{\nu_3} \vec{P}'$ and $\sigma_1 = \nu_1 \# \nu_2 \# \nu_3$. σ can obviously be rearranged as $M \rightarrow^{\nu_2} (\lambda x M_1 M'_2 \vec{P}) \rightarrow_\beta N = (M_1 \vec{P}) \rightarrow^{\nu_1 \# \nu_3} (M'_1 \vec{P}')$, which yields the result. \square

Lemma 4.12 1. Let $M = (\alpha \mu\beta M_1)$ and $N = M_1[\beta := \alpha]$. If N is strongly normalizable, then M is also strongly normalizable and $\eta(M) = \eta(N) + 1$.

2. Let $M = \mu\alpha(\alpha M_1)$ be a θ -redex. If M_1 is strongly normalizable, then M is also strongly normalizable and $\eta(M) = \eta(M_1) + 1$.

Proof

1. Assume σ is a reduction sequence starting from $(\alpha \mu\beta M_1)$. We prove $|\sigma| \leq \eta(N) + 1$, from which the result follows. Let $\sigma = [R]\#\sigma'$ for some σ' . We distinguish the various cases according to the form of σ .

- (a) If $(\alpha \mu\beta M_1) \rightarrow_\rho^R M_1[\beta := \alpha] \rightarrow^{\sigma'} M_2$, where $\sigma = [R]\#\sigma'$, then the result obviously follows.
- (b) If $(\alpha \mu\beta M_1) \rightarrow^R M_2 \rightarrow^{\sigma'} M_3$, where $M_2 \neq N$ and $\mu\beta M_1$ does not disappear in σ , then $M_3 = (\alpha \mu\beta M'_3)$ and $M_1 \rightarrow^\sigma M'_3$, which yields the result.
- (c) If $(\alpha \mu\beta M_1) \rightarrow^R M_2 \rightarrow^{\sigma'} M_3$, where $M_2 \neq N$ and $\mu\beta M_1$ disappears in σ . Then $(\alpha \mu\beta M_1) \rightarrow^{\sigma''} (\alpha \mu\beta(\beta M_k)) \rightarrow_\theta (\alpha M_k) \rightarrow^{\sigma'''} M_3$, where $\mu\beta M_1$ does not disappear in σ'' . We have $(\alpha \mu\beta M_1) \rightarrow_\rho M_1[\beta := \alpha] \rightarrow^{\sigma''[\beta := \alpha]} (\beta M_k)[\beta := \alpha] = (\alpha M_k) \rightarrow^{\sigma'''} M_3$, and the latter reduction sequence is equal in length to σ . By this the result follows.

The reverse direction is obvious.

2. Similar to the above one. □

Lemma 4.13 Let $M = (\mu\alpha M_1 M_2 \vec{P})$ and $N = (\mu\alpha M_1[\alpha :=_r M_2] \vec{P}')$.

- 1. If $\alpha \in Fv(M_1)$ and N is strongly normalizable, then M is also strongly normalizable $\eta(M) = \eta(N) + 1$.
- 2. If $\alpha \notin Fv(M_1)$ and N, M_2 are strongly normalizable, then M is also strongly normalizable and $\eta(M) = \eta(N) + \eta(M_2) + 1$.

Proof

1. Let $M \rightarrow^\sigma M^*$. We prove $|\sigma| \leq \eta(N) + 1$, from this $\eta(M) \leq \eta(N) + 1$ follows.

- (a) The redex $R = (\mu\alpha M_1 M_2)$ is involved in σ .
 - i. If $\mu\alpha M_1$ does not disappear in σ , $(\mu\alpha M_1 M_2 \vec{P}) \rightarrow^{\sigma'} (\mu\alpha M'_1 M'_2 \vec{P}') \rightarrow_\mu (\mu\alpha M'_1[\alpha :=_r M'_2] \vec{P}') \rightarrow^{\sigma''} M^*$. Then, since $\alpha \in Fv(M_1)$, by Lemmas 3.15 and 3.14, the reduction sequence $(\mu\alpha M_1 M_2 \vec{P}) \rightarrow^r (\mu\alpha M_1[\alpha :=_r M_2] \vec{P}) \rightarrow (\mu\alpha M_1[\alpha :=_r M'_2] \vec{P}) \rightarrow (\mu\alpha M'_1[\alpha :=_r M'_2] \vec{P}') \rightarrow^{\sigma''} M^*$ has length at least $|\sigma|$, by which the assertion follows.
 - ii. If $\mu\alpha M_1$ disappears in σ , $(\mu\alpha M_1 M_2 \vec{P}) \rightarrow^{\sigma'} (\mu\alpha(\alpha M'_1) M'_2 \vec{P}') \rightarrow_\theta (M'_1 M'_2 \vec{P}') \rightarrow^{\sigma''} M^*$. Then, since $\alpha \in Fv(M_1)$, by Lemmas 3.15 and 3.14, the sequence $(\mu\alpha M_1 M_2 \vec{P}) \rightarrow_\mu (\mu\alpha M_1[\alpha :=_r M_2] \vec{P}) \rightarrow (\mu\alpha(\alpha M'_1)[\alpha :=_r M'_2] \vec{P}') = (\mu\alpha(\alpha (M'_1 M'_2)) \vec{P}') \rightarrow_\theta (M'_1 M'_2 \vec{P}') \rightarrow^{\sigma''} M^*$ has length at least $|\sigma|$, which yields the result.
- (b) The redex $R = (\mu\alpha M_1 M_2)$ is not involved in σ .

- i. If $\mu\alpha M_1$ does not disappear in σ , that is, $(\mu\alpha M_1 M_2 \vec{P}) \rightarrow M^* = (\mu\alpha M'_1 M'_2 \vec{P}')$. Then, since $\alpha \in Fv(M_1)$, we can apply Lemmas 3.15 and 3.14 to assert that $(\mu\alpha M_1 M_2 \vec{P}) \rightarrow^R (\mu\alpha M_1[\alpha :=_r M_2] \vec{P}) \rightarrow (\mu\alpha M'_1[\alpha :=_r M'_2] \vec{P}')$ has length at least $|\sigma| + 1$.
- ii. If $\mu\alpha M_1$ disappears in σ , $(\mu\alpha M_1 M_2 \vec{P}) \rightarrow (\mu\alpha(\alpha M'_1) M'_2 \vec{P}') \rightarrow_\theta (M'_1 M'_2 \vec{P}') \rightarrow M^*$. By Lemmas 3.14 and 3.15, the sequence $(\mu\alpha M_1 M_2 \vec{P}) \rightarrow^R (\mu\alpha M_1[\alpha :=_r M_2] \vec{P}) \rightarrow (\mu\alpha(\alpha M'_1)[\alpha :=_r M'_2] \vec{P}') = (\mu\alpha(\alpha (M'_1 M'_2)) \vec{P}') \rightarrow_\theta (M'_1 M'_2 \vec{P}') \rightarrow M^*$ has length at least $|\sigma| + 1$, which proves the assertion.

The reverse direction is obvious.

2. The proof of $\eta(M) \leq \eta(N) + \eta(M_2) + 1$ is similar to the first part of the proof of Lemma 4.13. In this case the verification is made easier by the fact that, since $\alpha \notin Fv(M_1)$, $\mu\alpha M_1$ does not disappear in a reduction sequence starting from M . For the converse, let $N \rightarrow^\sigma N'$ and $M_2 \rightarrow^\nu M'_2$. Then $(\mu\alpha M_1 M_2 \vec{P}) \rightarrow^\nu (\mu\alpha M_1 M'_2 \vec{P}) \rightarrow_\mu (\mu\alpha M_1 \vec{P}) \rightarrow^\sigma N'$ is a reduction sequence starting from M , which means that $\eta(N) + \eta(M_2) + 1 \leq \eta(M)$. \square

4.3 The general case

In what follows we transform every $\lambda\mu$ -term M into a $\lambda\mu I$ -term $[M]_k$ with some $k \geq 0$ such that $\eta(M) \leq \eta([M]_k)$, by which, using Corollary 4.9, we can obtain a bound for $\eta(M)$ also.

At this point our presentation slightly differs from that of Xi ([27]). We have reformulated the translation in [27], hence we were able to avoid the minor mistake in [27] when computing the complexity of the obtained $\lambda\mu I$ -terms. For a detailed explanation see [3]. The interesting fact for Theorem 4.22, which is the main result of the paper, is, however, that we get the same bound for the simply typed $\lambda\mu$ -calculus as Xi obtained for the λ -calculus, *mutatis mutandis*. Namely, if we restrict the notion of the rank of a term in Definition 4.1 by taking into consideration the β -redex only, we get the result of Xi for the λ -calculus as a special case of Theorem 4.22.

Definition 4.14 1. Let $\mathcal{V} = \{v_{(A,B)} \mid A, B \text{ are types}\}$ be a set of distinguished variables such that for all A, B we have $v_{(A,B)} : A \rightarrow (B \rightarrow A)$, where $v_{(A,B)}$ are either constants or new variables. Let $M : A$ and $N : B$ be typed $\lambda\mu$ -terms. We denote the term $((v_{(A,B)} M) N)$ by $\langle M, N \rangle$.

2. Let M be a term and $k \geq 0$. The $\lambda\mu$ -term $[M]_k$ assigned to M is defined as follows.

- $[M]_k = M$, if M is a variable,
- $[M]_k = \lambda x \lambda y_1 \dots \lambda y_m \langle ([M_1]_k y_1 \dots y_m), x \rangle$, if $M = \lambda x M_1$ such that $lh(type(M)) \leq k$ and $type(M_1) = A_1 \rightarrow \dots \rightarrow A_m \rightarrow B$, $type(y_i) = A_i$ ($1 \leq i \leq m$) and B is atomic,
- $[M]_k = \lambda x \langle [M_1]_k, x \rangle$, if $M = \lambda x M_1$ and $lh(type(M)) > k$,
- $[M]_k = \mu\alpha \langle [M_1]_k, (\alpha z) \rangle$, if $M = \mu\alpha M_1$, where $\alpha \notin Fv(M_1)$ and z is a new variable such that $type(M) = type(z)$,
- $[M]_k = \mu\alpha [M_1]_k$, if $M = \mu\alpha M_1$ and $\alpha \in Fv(M_1)$,

- $[M]_k = (\alpha [M_1]_k)$, if $M = (\alpha M_1)$,
- $[M]_k = ([M_1]_k [M_2]_k)$, if $M = (M_1 M_2)$.

3. For each term M and each $k \geq 0$, we define the contexte $\Gamma_{M,k}$ wich contains the constants $v_{(A,B)}$ of $[M]_k$ with their type $A \rightarrow (B \rightarrow A)$.

Observe that in the definition above the translations for λ - and μ -abstractions differ. The underlying reason is the fact that in a β -reduction the λ -abstraction disappears while this is not the case concerning a μ -reduction. Hence, in order to ensure the validity of Lemma 4.21, we must make sure that the translation of a term with β -redex as head redex can be continued even after the reduction with the head redex. The main aim with the translation is to produce a $\lambda\mu$ I-term $[M]_k$ from M such that the relation $\eta(M) \leq \eta([M]_k)$ should be valid, which is the statement of Lemma 4.21. To achieve this, we reproduce the original M inside its translation $[M]_k$ in a sense, since, in general, the translation does not respect reduction, that is, if $M \rightarrow N$, then it is not necessarily the case that $[M]_k \rightarrow [N]_k$.

The next four lemmas describe some intuitively clear properties of the translation.

Lemma 4.15 *Let M be a term and $k \geq 0$.*

1. $[M]_k$ is a $\lambda\mu$ I-term.
2. $\alpha \in Fv(M)$ iff $\alpha \in Fv([M]_k)$ and if $x \in Fv(M)$, then $x \in Fv([M]_k)$.

Proof By induction on $comp(M)$. □

Observe that, in the case of λ -variables, $Fv(M) \subseteq Fv([M]_k)$, since M can contain free variables of the form $v_{(A,B)}$ besides its original parameters.

Lemma 4.16 *If M is a term and $k \geq 0$, then $rank([M]_k) = rank(M)$.*

Proof By induction on $comp(M)$. □

Lemma 4.17 *If M be a term and $k \geq 0$, then $comp([M]_k) \leq (2k + 3).comp(M)$.*

Proof The only nontrivial case is $M = \lambda x M_1$. Let $lh(type(\lambda x M_1)) = l$. If $k < l$, then $comp([M]_k) = comp(\lambda x ([M_1]_k, x)) = comp([M_1]_k) + 3 \leq (2k + 3).comp(M)$. If $k \geq l$, then, for some $m \leq l$, we obtain by the induction hypothesis $comp([M]_k) = comp(\lambda x \lambda y_1 \dots \lambda y_m \langle ([M_1]_k y_1 \dots y_m), x \rangle) = comp([M_1]_k) + 2m + 3 \leq (2k + 3).comp(M)$. □

Lemma 4.18 *Let M, N be terms and $k \geq 0$.*

1. If $\Gamma \vdash M : A$, then $\Gamma, \Gamma_{M,k} \vdash [M]_k : A$.
2. $[M]_k[x := [N]_k] = [M[x := N]]_k$,
3. $[M]_k[\alpha :=_r [N]_k] = [M[\alpha :=_r N]]_k$,
4. $[M]_k[\beta := \alpha] = [M[\beta := \alpha]]_k$.

Proof By induction on $comp(M)$. □

Our aim is to prove $\eta(M) \leq \eta([M]_k)$. The assertions of Subsection 4.2 and Lemma 4.19 prepare the proof of that statement, which is the claim of Lemma 4.21.

Lemma 4.19 *If $M \rightarrow M'$ and $\text{rank}(M) \leq k$, then $\eta([M']_k) + 1 \leq \eta([M]_k)$.*

Proof By induction on $\text{comp}(M)$.

1. If $M = \lambda x M_1$, the induction hypothesis applies.
2. If $M = (\lambda x M_1 M_2 \dots M_n)$. We have $lh(\text{type}(\lambda x M_1)) \leq k$ by virtue of the assumption $\text{rank}(M) \leq k$. Let $\text{type}(M_1) = A_1 \rightarrow \dots A_m \rightarrow B$, where B is atomic. Let $M' = (M_1[x := M_2] \dots M_n)$, otherwise the induction hypothesis applies. Since B is atomic, $m \geq n - 2$ holds. Then $[M]_k \rightarrow \lambda y_1 \dots \lambda y_m \langle ([M_1]_k[x := [M_2]_k] y_1 \dots y_m), [M_2]_k \rangle \dots [M_n]_k \twoheadrightarrow \lambda y_{n-1} \dots \lambda y_m \langle ([M_1]_k[x := [M_2]_k] \dots [M_n]_k y_{n-1} \dots y_m), [M_2]_k \rangle$. Lemma 4.18 gives $([M_1]_k[x := [M_2]_k] [M_3]_k \dots [M_n]_k) = ([M_1[x := M_2]]_k [M_3]_k \dots [M_n]_k) = ([M_1[x := M_2] M_3 \dots M_n]_k) = [M']_k$, by which the result follows.
3. If $M = (\mu \alpha M_1 M_2 \dots M_n)$, we may assume again that $M \rightarrow^R M'$, where $R = (\mu \alpha M_1 M_2)$.
 - If $\alpha \in Fv(M_1)$, let $M' = (\mu \alpha M_1[\alpha :=_r M_2] \dots M_n)$. We have, by Lemma 4.18, $[M]_k = (\mu \alpha [M_1]_k [M_2]_k \dots [M_n]_k) \rightarrow (\mu \alpha [M_1]_k[\alpha :=_r [M_2]_k] \dots [M_n]_k) = (\mu \alpha [M_1[\alpha :=_r M_2]]_k \dots [M_n]_k) = [M']_k$.
 - If $\alpha \notin Fv(M_1)$, then $M' = \mu \alpha M_1 M_3 \dots M_n$ and $[M]_k \rightarrow (\mu \alpha \langle [M_1]_k, (\alpha(z [M_2]_k)) \rangle [M_3]_k \dots [M_n]_k)$. We may assume $\alpha \notin Fv(M_i)$ ($1 \leq i \leq k$). Then Lemma 4.10 gives $\eta((\mu \alpha \langle [M_1]_k, (\alpha(z [M_2]_k)) \rangle [M_3]_k \dots [M_n]_k)) \geq \eta((\mu \alpha \langle [M_1]_k, (\alpha z) \rangle [M_3]_k \dots [M_n]_k)) + \eta([M_2]_k) + 1$. Moreover, by induction on n , we obtain that $\eta([M']_k) \leq \eta((\mu \alpha \langle [M_1]_k, (\alpha z) \rangle [M_3]_k \dots [M_n]_k))$, by which the result follows.
4. If $M = (\alpha M_1)$, the only interesting case is $M = (\alpha \mu \beta M'_1) \rightarrow M'_1[\beta := \alpha]$. If $\beta \in Fv(M'_1)$, then $[M]_k = (\alpha \mu \beta [M'_1]_k)$. Otherwise, $[M]_k = (\alpha \mu \beta \langle [M'_1]_k, (\beta z) \rangle)$. Applying Lemma 4.18, in both cases we obtain the result.
5. The case $M = \mu \alpha M_1$ is analogous to the previous one.
6. If $M = (x M_1 \overrightarrow{P})$, the induction hypothesis applies.

□

Prior to proving the next lemma, we demonstrate with an example that the hypothesis $\text{rank}(M) \leq k$ was indeed necessary for the validity of Lemma 4.19.

Example 4.20 *Let $M = ((\lambda x \lambda y y x) y)$. Then $M' = (\lambda y y y)$. Assume $x, y : A$. Then $\text{rank}(\lambda x \lambda y y) = 2$, which means $\text{rank}(M) = 2$. Let $k = 1$. Then $[M]_1 = ((\lambda x \langle [\lambda y y]_1, x \rangle x) y)$, and $[M']_1 = (\lambda y \langle y, y \rangle y)$. Since $[\lambda y y]_1 = \lambda y \langle y, y \rangle$ is not a redex, we have $\eta([M]_1) = \eta([M']_1) = 1$, thus the statement of Lemma 4.19 is not valid for M .*

Lemma 4.21 *If M is a $\lambda\mu$ -term such that $\text{rank}(M) \leq k$, then $\eta(M) \leq \eta([M]_k)$.*

Proof By induction on $(\eta([M]_k), \text{comp}(M))$.

1. If $M = \lambda x M_1$, then, by the induction hypothesis we have the result.
2. If $M = (x M_1 \dots M_n)$, then, by the induction hypothesis, $\eta(M) = \eta(M_1) + \dots + \eta(M_n) \leq \eta([M_1]_k) + \dots + \eta([M_n]_k) = \eta([M]_k)$.

3. If $M = (\lambda x M_1 M_2 \dots M_n)$, let $M' = (M_1[x := M_2] \dots M_n)$. It follows from Lemma 4.2 that $\text{rank}(M') \leq k$.
 - $x \in Fv(M_1)$: By Lemmas 4.19 and 4.11 and the induction hypothesis, $\eta(M) = \eta(M') + 1 \leq \eta([M']_k) + 1 \leq \eta([M]_k)$.
 - $x \notin Fv(M_1)$: $[M]_k \rightarrow_\beta \lambda y_1 \dots \lambda y_m \langle ([M_1]_k y_1 \dots y_m), [M_2]_k \rangle \dots [M_n]_k = U$. By Lemma 4.11, we are ready, if we prove $\eta([(M_1 M_3 \dots M_n)]_k) \leq \eta(U)$. By the choice of m , we have $m \geq n - 2$, hence $U \rightarrow \langle ([M_1]_k [M_3]_k \dots [M_m]_k), [M_2]_k \rangle \dots [M_n]_k$, from which the conclusion follows.
4. Let $M = (\mu \alpha M_1 M_2 \dots M_n)$.
 - If $\alpha \in Fv(M_1)$, let $M' = (\mu \alpha M_1[\alpha :=_r M_2] \dots M_n)$. Then $\text{rank}(M') \leq k$ by Lemma 4.2 again. We have, by Lemmas 4.19, 4.13 and the induction hypothesis, $\eta(M) = \eta(M') + 1 \leq \eta([M']_k) + 1 = \eta([M]_k)$.
 - If $\alpha \notin Fv(M_1)$, let $M' = (\mu \alpha M_1 M_3 \dots M_n)$. We have $[M]_k \rightarrow (\mu \alpha \langle [M_1]_k, (\alpha (z [M_2]_k)) \rangle [M_3]_k \dots [M_n]_k)$, which, together with Lemmas 4.13, 4.10, 4.2 and the induction hypothesis, yields that $\eta(M) \leq \eta(M') + \eta(M_2) + 1 \leq \eta([M']_k) + \eta([M_2]_k) + 1 \leq \eta([M]_k)$.
5. Let $M = \mu \alpha M_1$.
 - Assume $\alpha \in Fv(M_1)$. If $\mu \alpha M_1 = \mu \alpha (\alpha M_2)$ is a θ -redex, then, by Lemmas 4.12, 4.2 and the induction hypothesis, $\eta(M) = \eta(M_2) + 1 \leq \eta([M_2]_k) + 1 = \eta([M]_k)$. Otherwise, let $\mu \alpha M_1 \rightarrow M'$. Since $\mu \alpha M_1$ is not a θ -redex, we have $M' = \mu \alpha M'_1$ together with $\text{rank}(M') \leq k$. By Lemma 4.19, we can apply the induction hypothesis to M' , that is, $\eta(\mu \alpha M'_1) + 1 \leq \eta([\mu \alpha M'_1]_k) + 1 \leq \eta([\mu \alpha M_1]_k)$. But M' was arbitrary and $\eta(M) = \max\{\eta(M') + 1 \mid M \rightarrow M'\}$, which proves our assertion.
 - If $\alpha \notin Fv(M_1)$, then we can apply the induction hypothesis to M_1 .
6. Let $M = (\alpha \mu \beta M')$. Similar to the previous case by using Lemma 4.12. □

Theorem 4.22 *If M is a $\lambda\mu$ -term such that $\text{rank}(M) = k$, then every $\beta\mu\rho\theta$ -reduction sequence starting from M is of length less than $\text{tower}(k + 1, (2k + 3) \cdot \text{comp}(M))$.*

Proof We obtain, by Lemma 4.17, $\text{comp}([M]_k) \leq (2k + 3) \cdot \text{comp}(M)$ and, by Lemma 4.16, $\text{rank}([M]_k) = \text{rank}(M)$. These, together with Corollary 4.9 and Lemma 4.21, imply $\eta(M) \leq \eta([M]_k) < \text{tower}(k + 1, \text{comp}([M]_k)) \leq \text{tower}(k + 1, (2k + 3) \cdot \text{comp}(M))$. □

5 Concluding remarks

In what follows, we give a short account of other possibilities for obtaining bounds for the reduction sequences in the $\lambda\mu$ -calculus. We could resort to the idea of translating the $\lambda\mu$ -calculus in the λ -calculus such that the sizes of the translated terms and the lengths of their reduction sequences would depend on the sizes and lengths of the original ones in an estimated way and then we could apply Xi's method to extract a bound for the lengths of reduction sequences in the $\lambda\mu$ -calculus, as well. By examining this idea, we find that we were not able to simulate every reduction rule, if we apply the already existing translations, and even the bound would be much worse than the one appearing in our result. We investigate these questions in detail below.

5.1 A possible attempt to calculate an upper bound for the $\lambda\mu\rho$ -calculus

In the following observations we confine our attention to the case of the $\lambda\mu\rho$ -calculus. For establishing a bound for the lengths of reduction sequences of the $\lambda\mu\rho$ -calculus it seems to be a natural idea to try to transform a reduction sequence of the $\lambda\mu\rho$ -calculus into a reduction sequence of the λ -calculus. We go round this approach a little bit more detailed: we present the CPS-translation from the simply-typed $\lambda\mu\rho$ -calculus to the simply-typed λ -calculus introduced by de Groote [14], and then we give an account of the possibilities of finding an appropriate bound with this method. The notation for the CPS-translation is taken from de Groote [14]. As to a bound for the simply-typed λ -calculus we regard the one presented in Xi [27].

Definition 5.1 *Let o be some distinguished atomic type.*

1. *For every type A , we define the three types A^o , $\sim A$ and \overline{A} by : $\sim A = A \rightarrow o$, $\overline{A} = \sim \sim A^o$, $\perp^o = o$, $X^o = X$, if X is atomic, and $(B \rightarrow C)^o = \overline{B} \rightarrow \overline{C}$.*
2. *Let Γ_λ (resp. Γ_μ) denote a λ -context (resp. μ -context), that is, a finite (possibly empty) set of declarations $x_1 : A_1, \dots, x_n : A_n$ (resp. $\alpha_1 : \neg B_1, \dots, \alpha_m : \neg B_m$). We define $\overline{\Gamma}_\lambda$ (resp. $\sim \Gamma_\mu^o$) by $x_1 : \overline{A_1}, \dots, x_n : \overline{A_n}$ (resp. $\alpha_1 : \sim B_1^o, \dots, \alpha_m : \sim B_m^o$).*

We suppose that the μ -variables of the $\lambda\mu\rho$ -calculus are also λ -variables of the λ -calculus.

Definition 5.2 *The CPS-translation \overline{M} of a $\lambda\mu\rho$ -term M is defined as follows.*

- $\overline{x} = \lambda k(x \ k)$,
- $\overline{\lambda x M} = \lambda k(k \ \lambda x \overline{M})$,
- $\overline{(M \ N)} = \lambda k(\overline{M} \ (\lambda m(m \ \overline{N} \ k)))$,
- $\overline{\mu \alpha M} = \lambda \alpha(\overline{M} \ (\lambda k k))$,
- $\overline{(\alpha \ M)} = \lambda k(\overline{M} \ \alpha)$.

Lemma 5.3 *Let $M : A$ be a typable term with λ -context Γ_λ and μ -context Γ_μ . Then its CPS-translation, \overline{M} , is typable with contexts $\overline{\Gamma}_\lambda$ and $\sim \Gamma_\mu^o$.*

Definition 5.4 *Let $=_\lambda$ (resp. $=_\mu$) denote the relation defined as the reflexive, symmetric, transitive closure of the β -reduction (resp. that of the union of the β -, μ - and ρ -reductions). As usual, we consider terms differing in renaming of bound variables as equals.*

Then, in [15], de Groote proves the following result.

Lemma 5.5 *Let M, N be $\lambda\mu\rho$ -terms. Then $M =_\mu N$ iff $\overline{M} =_\lambda \overline{N}$.*

Unfortunately, in Lemma 5.5, $\overline{M} \rightarrow_\lambda \overline{N}$ does not hold generally, even if $M \rightarrow_\mu N$. So on one hand we cannot use the CPS-translation to imitate the reduction sequences in the $\lambda\mu\rho$ -calculus by reduction sequences in the λ -calculus. On the other hand there can be another drawback of this approach.

In general, we could make use of the CPS-translation for estimating bounds of reduction sequences if for any $M \rightarrow_\mu^\sigma NF(M)$ we could find a ν with $\overline{M} \rightarrow_\lambda^\nu NF(\overline{M})$ such that $|\sigma| \leq c \cdot |\nu|$ with some constant c , where $NF(M)$ and $NF(\overline{M})$ denote the (unique) normal form of M in the $\lambda\mu\rho$ -calculus and of \overline{M} in the λ -calculus,

respectively. In fact, we even know that $NF(\overline{M}) = \overline{NF(\overline{M})}$, where \overline{M} stands for the so called modified CPS-translation of the term M (de Groote [15]).

For the moment suppose for every reduction sequence $M \rightarrow^\sigma NF(M)$ we can find a reduction sequence ν such that $\overline{M} \rightarrow_\lambda^\nu NF(\overline{M})$ with $|\sigma| \leq c \cdot |\nu|$. By the result for the β -normalization in Xi [27], we would have for any ν as above $|\nu| < c \cdot \text{tower}(\text{rank}(\overline{M}) + 1, (2 \cdot \text{rank}(\overline{M}) + 3) \cdot \text{comp}(\overline{M}))$. On the other hand we have the following estimations.

Lemma 5.6 *Let M be a $\lambda\mu$ -term. Then $\text{rank}(\overline{M}) = 3 \cdot \text{rank}(M)$ and $2 \cdot \text{comp}(M) < \text{comp}(\overline{M})$.*

This means that the best estimation for the lengths of the reductions with this method would be greater than $c \cdot \text{tower}(3 \cdot \text{rank}(M) + 1, (12 \cdot \text{rank}(M) + 6) \cdot \text{comp}(M))$, and by the direct method this upper bound is $\text{tower}(\text{rank}(M) + 1, (2 \cdot \text{rank}(M) + 3) \cdot \text{comp}(M))$. At present, no CPS-translation which could yield a significantly better estimation is known to the authors.

5.2 A translation of the $\lambda\mu$ -calculus into the λ_c^* -calculus

Some years ago a new translation of the $\lambda\mu$ -calculus into a version of the λ -calculus formulated with recursive equations for types was discovered by David and Nour ([10]). This is somewhat simpler than the CPS-translation and provides an easy method of finding an estimation for the lengths of reduction sequences in the $\lambda\mu$ -calculus. We present a version of it, establishing a connection between the $\lambda\mu$ -calculus and a variant of the λ -calculus enlarged with some constants. The method traces back to Krivine ([16] and [17]), where he supplemented the typed-calculus with a constant of type $\forall X(\neg\neg X \rightarrow X)$.

Definition 5.7 *Enhance the set of types of the simply typed λ -calculus with an element \perp and define $\neg A$ as $A \rightarrow \perp$. Let X be an atomic type, add for each X a new constant c_X of type $\neg\neg X \rightarrow X$. Let us call the new calculus as λ_c^* . We define for each type A a closed λ_c^* -term T_A such that T_A has the type $\neg\neg A \rightarrow A$.*

- $T_\perp = \lambda y(y \ I)$, where $I = \lambda x x$,
- $T_X = c_X$, where X is atomic,
- $T_{A \rightarrow B} = \lambda x \lambda y(T_B \ \lambda z(x \ \lambda t(z \ (t \ y))))$.

We suppose that the μ -variables of the $\lambda\mu$ -calculus are also λ -variables of the λ -calculus.

Definition 5.8 *Let $k \geq 0$. We define a k -translation $|\cdot|_k$ of the set of $\lambda\mu$ -terms into the set of terms of the λ_c^* -calculus as follows.*

- $|x|_k = x$,
- $|\lambda x M|_k = \lambda x |M|_k$,
- $|(M \ N)|_k = (|M|_k \ |N|_k)$,
- $|\mu\alpha M|_k = (T_A \ \lambda\alpha |M|_k)$, if α has type $\neg A$ and $lh(A) \leq k$,
- $|\mu\alpha M|_k = (z \ |M|_k)$, if α has type $\neg A$ and $lh(A) > k$ and where $z : \perp \rightarrow A$ is a new variable,
- $|(\alpha \ M)|_k = (\alpha \ |M|_k)$.

In the above definition the μ -variables and its translated counterparts were denoted with the same letters. Let $\vdash_{\lambda\mu}$ and $\vdash_{\lambda_c^*}$ denote the typing relations in the $\lambda\mu$ - and in the λ^* -calculus, respectively. We have the following assertions.

Lemma 5.9 *Let $k \geq 0$ and M a $\lambda\mu$ -term. If $\Gamma \vdash_{\lambda\mu} M : A$, then $\Gamma \vdash_{\lambda_c^*} |M|_k : A$.*

Proof Straightforward. \square

Lemma 5.10 *Let $k \geq 0$, M, N be $\lambda\mu$ -terms such that $\text{rank}(M) \leq k$.*

If $M \rightarrow_{\lambda\mu} N$, then $|M|_k \rightarrow_{\lambda}^+ |N|_k$.

Proof Obviously, it is enough to check the relation $|(\mu\alpha M_1 M_2)|_k \rightarrow_{\lambda}^+ |\mu\alpha M_1[\alpha :=_r M_2]|_k$, where, necessarily, $k \geq lh(A)$ provided $\text{type}(\alpha) = \neg A$. \square

Lemma 5.11 *Let $k \geq 0$, M, N be $\lambda\mu$ -terms such that $\text{rank}(M) \leq k$.*

If $M \rightarrow^n N$, then $|M|_k \rightarrow^m |N|_k$ for some $m \geq n$.

Proof Follows from Lemmas 4.2 and 5.10. \square

Since no reduction rules are added to λ when defining λ_c^* , the method of Xi [27] for estimating the lengths of reduction sequences is also applicable to λ_c^* without any changes. We state without proof the following theorem.

Theorem 5.12 *Let M be a λ_c^* -term such that $\text{rank}(M) = k$. Then every reduction sequence starting from M has length less than $\text{tower}(k+1, (2k+3).\text{comp}(M))$.*

In order to establish a bound for the lengths of $\lambda\mu$ -reduction sequences we have to estimate the size of the translated terms as well.

Lemma 5.13 *If A is a type, then $\text{comp}(T_A) \geq 8.lh(A) + 3$.*

Proof Obvious. \square

Lemma 5.14 *If M is a $\lambda\mu$ -term such that $\text{rank}(M) = k$, then $\text{comp}(|M|_k) \leq (8k+4).\text{comp}(M)$.*

Proof By induction on $\text{comp}(M)$. We only check one of the cases. Let $M = (\mu\alpha M_1 M_2)$. Assume $\text{type}(\alpha) = \neg A$. Then, since $k \geq lh(A)$, we have, by Lemma 5.13 and the induction hypothesis, $\text{comp}(|M|_k) = \text{comp}(|\mu\alpha M_1|_k |M_2|_k) = \text{comp}((T_A \lambda\alpha |M_1|_k) + \text{comp}(|M_2|_k)) \leq (8k+4) + \text{comp}(|M_1|_k) + \text{comp}(|M_2|_k) \leq (8k+4).\text{comp}(M)$. \square

Theorem 5.15 *Let M be a $\lambda\mu$ -term such that $\text{rank}(M) = k$. Then every reduction sequence starting from M has length less than $\text{tower}(k+1, (2k+3)(8k+4).\text{comp}(M))$.*

Proof Follows from Theorem 5.12 and Lemma 5.14. \square

This method, however, is not applicable to the $\lambda\mu\rho\theta$ -calculus, since, in the cases of the ρ - and θ -reductions, Lemma 5.10 is not valid.

6 Future work

In this paper, we have shown how to find a bound for the lengths of reduction sequences in the simply typed $\lambda\mu$ -calculus. The bound depends only on the size of the term and on the maximum of the ranks of its redexes. The proof is accomplished by finding a bound for the $\lambda\mu$ I-terms first, then we apply a translation of an arbitrary term to a $\lambda\mu$ I-term, such that the lengths of reduction sequences do not decrease. We have also used a standardization result for the $\lambda\mu$ -calculus, the formulation of which was not entirely straightforward because of the presence of the ρ - and θ -rules. Our future work can be the following.

1. **Finding a term realizing the bound.** In the literature usually different bounds are found for the reductions of the λ -calculus. The question naturally arises, which bounds could be the closest ones? Could we amend the present bound considerably?
2. **Commutation lemmas for the ρ - and θ -rules.** If we considered only the β - and μ -rules, our proof would simplify considerably, especially those concerning standardization. However, treating the other rules, the ρ - and the θ -rules may not be easy, since we had to prove by commutation lemmas that they could be postponed together with maintaining an upper bound for the lengths of the reductions. It would be good to see whether this approach simplifies the presentation of our results or not.
3. **Other rules for the $\lambda\mu$ -calculus.** The $\lambda\mu$ -calculus has other kinds of reductions. For example, one can prohibit two consecutive μ -abstractions ($\mu\alpha\mu\beta M$) or μ -variable applications ($((\alpha (\beta M)))$) ([18]). Parigot has also proposed a rule which prohibits that a λ -abstraction should immediately follow a μ -variable ([24]). Another rule is also worth considering: $(N \mu\alpha M) \rightarrow_{\mu'} \mu\alpha M[\alpha : =_l N]$, where $M[\alpha : =_l N]$ is obtained from M by replacing every subterm in M of the form (αU) by $(\alpha (N U))$. This rule is the symmetric counterpart of the μ -rule, the addition of which makes the $\lambda\mu$ -calculus non-confluent, but the strong normalization still holds ([8], [9]). It would also be interesting to try to find bounds with these rules.
4. **Other classical calculi.** It would be interesting to find an upper bound for the lengths of the reductions in other classical calculi ([2], [4]). Are the methods presented here applicable for them, too?

Acknowledgements. We wish to thank René David for helpful discussions.

References

- [1] H. P. Barendregt. *The Lambda Calculus: Its Syntax and Semantics*, Studies in Logic and the Foundations of Mathematics, Volume 103, North Holland, 1985.
- [2] F. Barbanera and S. Berardi. *A symmetric lambda calculus for classical program extraction*, Information and computation (125), pp. 103-117, 1996.
- [3] P. Battyányi. *Normalization properties of symmetric logical calculi*, PhD thesis, University of Chambéry, 2007.
- [4] P.-L. Curien and H. Herbelin. *The duality of computation*, ACM SIGPLAN Notices, 2000.
- [5] R. David. *Normalization without reducibility*, Annals of Pure and Applied Logic (107), pp. 121-130, 2001.
- [6] R. David and K. Nour. *A short proof of the strong normalization of the simply typed $\lambda\mu$ -calculus*, Schedae Informaticae (12), 2003.
- [7] R. David and K. Nour. *Why the usual candidates of reducibility do not work for the symmetric $\lambda\mu$ -calculus*, Electronic Notes in Theoretical Computer Science (140), 2005.
- [8] R. David and K. Nour. *Arithmetical proofs of strong normalization results for the symmetric $\lambda\mu\mu'$ -calculus*, TLCA 2005, Lecture Notes in Computer Science (3461), pp. 162-178, Springer Verlag, Berlin, 2005.

- [9] R. David and K. Nour. *Arithmetical proofs of strong normalization results for symmetric lambda calculi*, Fundamenta Informaticae (77), pp. 1001-1022, 2007.
- [10] R. David and K. Nour. *An arithmetical proof of the strong normalization for the λ -calculus with recursive types*, TLCA 2007, Lecture Notes in Computer Science (4583), Springer Verlag, Berlin, 2007.
- [11] T. Griffin. *A formulae-as-type notion of control*, POPL 1990, pp. 47-58, ACM Press, New York, 1990.
- [12] P. de Groote. *An environment machine for the $\lambda\mu$ -calculus*, Mathematical Structure in Computer Science (8), 1998.
- [13] C. R. Murthy. *An evaluation semantics for classical proofs*, Proceedings of the sixth annual IEEE symposium, pp. 96-107, 1991.
- [14] P. de Groote. *Strong normalization of classical natural deduction with disjunction*, TLCA 2001, Lecture Notes in Computer Science (2044), Springer Verlag, Berlin, 2001.
- [15] P. de Groote. *A CPS-translation of the $\lambda\mu$ -calculus*, Lecture Notes in Computer Science (787), Springer Verlag, Berlin, 1994.
- [16] J.-L. Krivine. *Lambda-calcul, évaluation paresseuse et mise en mémoire*, Theoretical Informatics and Applications. Vol. 25,1 pp. 67-84, 1991.
- [17] J.-L. Krivine. *Classical logic, storage operators and 2nd order lambda-calculus*, Annals of Pure and Applied Logic (68), pp. 53-78, 1994.
- [18] K. Nour. *La valeur d'un entier classique en $\lambda\mu$ -calcul*, Archive for Mathematical Logic (36), pp. 461-471, 1997.
- [19] M. Parigot. *$\lambda\mu$ -calculus: an algorithmic interpretation of classical natural deduction*, Lecture Notes in Computer Science (624), pp. 190-201, Springer Verlag, Berlin, 1992.
- [20] M. Parigot. *Classical proofs as programs*, Lecture Notes in Computer Science (713), pp. 263-276, Springer Verlag, Berlin, 1993.
- [21] M. Parigot. *Proofs of strong normalization for second order classical natural deduction*, Journal of Symbolic Logic (62), pp. 1461-1479, 1997.
- [22] M. Parigot. *Free Deduction: An Analysis of "Computations" in Classical Logic*, Lecture Notes in Computer Science (592), pp. 361-380, Springer-Verlag, Berlin, 1990.
- [23] N. J. Rehof and M. H. Sørensen. *The λ_{Δ} -calculus*, Lecture Notes in Computer Science (789), pp. 516-542, 1994.
- [24] W. Py. *Confluence en $\lambda\mu$ -calcul*, PhD thesis, University of Chambéry, 1998.
- [25] P. Rozière. *Déduction naturelle classique et $\lambda\mu$ -calcul*, Course notes, 2003.
- [26] W. W. Tait. *Intensional Interpretations of Functionals of Finite Type I*, Journal of Symbolic Logic (32), pp. 198-212, 1967.
- [27] H. Xi. *Upper bounds for standardisation and an application*, Journal of Symbolic Logic (64), pp. 291-303, 1999.
- [28] H. Xi. *Development separation in lambda-calculus*, WOLLIC 2005, Electronic Notes in Theoretical Computer Science (143), pp. 207-221, 2005.

A Appendix: Some technical definitions

In this appendix the notions describing the effects of performing redexes in a term are defined. We give formal definitions of residuals, that is, subterms remaining after executing a redex, we provide a precise notion of a redex being involved in a reduction sequence, as well. The treatment presented here is not heavily exploited in the paper, though all the proofs above could be reformulated in a way to fit the presentation of this section. In order to make the treatment precise, we resort to the technique of identifying subterms by finite sequences of indexes. First we settle how to identify a subterm by its index, this is achieved by the operator *occ*. Then we are able to describe precisely the index set of the residuals created in the term by executing a redex. This is accomplished with the function *add*. By now we have reached the point, where the actual set of residuals can be defined with respect to a sequence of reductions. This is done by the operator *des*, and, at this stage, we are able to define involvement of a redex in relation to a reduction sequence.

Definition A.1 Let $\Pi = (\{\lambda x \mid x \in \mathcal{V}\} \cup \{l_\alpha, r_\alpha \mid \alpha \in \mathcal{W}\} \cup \{\mu\alpha \mid \alpha \in \mathcal{W}\} \cup \{l, r\})^{<\infty}$. The elements of Π are called *indexes*.

1. A *path* (or *address*) for terms is a finite sequence consisting of elements of Π , that is, an element of the set $\text{Paths} = \Pi^{<\infty}$. The empty path will be denoted by $[]$.
2. Let $\xi \# \xi'$ denote the concatenation of $\xi, \xi' \in \text{Paths}$.
3. If $\xi \in \text{Paths}$ and $p \in \Pi$, then $[p :: \xi]$ is the concatenation of $[p]$ and ξ . Then we have $[p :: \xi] = [p] \# \xi$.
4. If $\xi \in \text{Paths}$, $\text{lst}(\xi)$ denote the last element of ξ provided $\xi \neq []$, undefined otherwise.
5. Let $\sigma = [p_1, \dots, p_n]$ ($n \geq 1$) be a path (or address). We say that σ' is an *initial subsequent* or *subpath* of σ (denoted by $\sigma' \leq \sigma$), if $\sigma' = [p_1, \dots, p_i]$ for some $1 \leq i \leq n$. The notation $\sigma_1 \approx \sigma_2$ indicates the fact that neither $\sigma_1 \leq \sigma_2$ nor $\sigma_2 \leq \sigma_1$ holds.

Definition A.2 1. The function *occ* is such that if M is a term and π is a path, then $\text{occ}(M, \pi)$ is the subterm of M at π provided it is defined.

- $\text{occ}(M, []) = M$,
 - $\text{occ}((M_1 M_2), [l :: s]) = \text{occ}(M_1, s)$, if $M_1 \notin \mathcal{W}$,
 - $\text{occ}((M_1 M_2), [r :: s]) = \text{occ}(M_2, s)$, if $M_1 \notin \mathcal{W}$,
 - $\text{occ}(\lambda x M, [\lambda x :: s]) = \text{occ}(M, s)$,
 - $\text{occ}((\alpha M), [l_\alpha]) = \alpha$,
 - $\text{occ}((\alpha M), [r_\alpha :: s]) = \text{occ}(M, s)$,
 - $\text{occ}(\mu\alpha M, [\mu\alpha :: s]) = \text{occ}(M, s)$.
2. The expression $\text{occ}(M, \pi)$ is undefined if it is not in one of the forms appearing in the left-hand sides of the above equations. We say that π is an *occurrence* of N in M , if $\text{occ}(M, \pi) = N$. The notation $\xi \in M$ indicates the fact that $\text{occ}(M, \xi)$ is defined.
 3. If N is a subterm of M , we write $N \leq M$. We can observe that $\pi_2 \leq \pi_1$ iff $\text{occ}(M, \pi_1) \leq \text{occ}(M, \pi_2)$. We denote by $\text{Sbt}(M)$ the set of subterms of a term M . If S is an arbitrary set of terms, let $\mathcal{P}(S)$ stand for the set of its subsets.

4. A path ξ is said to be an occurrence of a μ -variable α (resp. λ -variable x) if $\text{occ}(M, \xi) = \alpha$ (resp. $\text{occ}(M, \xi) = x$).

Example A.3 Let $M = (\lambda x \mu \alpha (x (\alpha y)) (\alpha x))$. Then the occurrences of x in M are $[l, \lambda x, \mu \alpha, l]$ and $[r, r_\alpha]$, while the occurrences of α are $[l, \lambda x, \mu \alpha, r, l_\alpha]$ and $[r, l_\alpha]$.

Definition A.4 1. Let M be a term. The function $\text{bn} : \mathcal{T} \times \text{Paths} \rightarrow \mathcal{T}$ defined below has the following property: if $\text{occ}(M, \xi) = x$ (resp. $\text{occ}(M, \xi) = \alpha$), then $\text{bn}(M, \xi)$ is the smallest subterm of M containing $\text{occ}(M, \xi)$ in which the variable x (resp. the variable α) becomes bound.

- If $\text{occ}(M, \xi) = \alpha$, then let $\text{bn}(M, \xi) = \text{occ}(M, \zeta)$, where ζ is the longest initial segment of ξ such that $\text{occ}(M, \zeta) = \mu \alpha Q$ if such a Q exists.
 - If $\text{occ}(M, \xi) = x$, then $\text{bn}(M, \xi) = \text{occ}(M, \zeta)$, where ζ is the longest initial segment of ξ such that $\text{occ}(M, \zeta) = \lambda x Q$ if such a Q exists.
 - Let $\text{bn}(M, \xi)$ be undefined otherwise.
2. An occurrence ξ of α (resp. x) is free in M if $\text{bn}(M, \xi)$ is undefined. If M is a term $Fv_\mu(M)$ and $Fv_\lambda(M)$ denote the set of free μ - and λ -variables of M , respectively.

Example A.5 Let $M = (\lambda x \mu \alpha (x (\alpha y)) (\alpha x))$, $\xi_1 = [l, \lambda x, \mu \alpha, l]$ and $\xi_2 = [l, \lambda x, \mu \alpha, r, l_\alpha]$. Then the first occurrences of x in M , $\text{occ}(M, \xi_1) = x$, is bound, since ξ_1 contains λx , but the second one, $\text{occ}(M, [r, r]) = x$, is free. Similarly for α , the first occurrence of α , with address ξ_2 , is bound, since the index sequence belonging to that occurrence contains a $\mu \alpha$, and the second occurrence, $[r, l_\alpha]$, is free. We have $\text{bn}(M, \xi_1) = \lambda x \mu \alpha (x (\alpha y))$ and $\text{bn}(M, \xi_2) = \mu \alpha (x (\alpha y))$.

Definition A.6 Let M be a term and $\xi = [a_1, \dots, a_n] \in \text{Paths}$. Assume $\text{occ}(M, \xi)$ is defined. We denote by $l(M, \alpha, \xi)$ the address obtained from ξ by inserting simultaneously an element l after each a_i for which $a_i = r_\alpha$ and $[\xi_{i-1} :: l_\alpha]$ is a free occurrence of α in M , where $\xi_{i-1} = [a_1, \dots, a_{i-1}]$ and $\xi_0 = []$. We write briefly $l(\xi)$ if M and α are clear from the context.

Example A.7 Let $M = (\alpha (x \mu \alpha (\alpha y)) (\alpha \lambda x y))$, $\xi_1 = [l, r_\alpha, l]$, $\xi_2 = [l, r_\alpha, r, \mu \alpha, r_\alpha]$ and $\xi_3 = [r, r_\alpha]$. Then $l(M, \xi_1, \alpha) = [l, r_\alpha, l, l]$, $l(M, \xi_2, \alpha) = [l, r_\alpha, l, r, \mu \alpha, r_\alpha]$ and $l(M, \xi_3, \alpha) = [r, r_\alpha, l]$, respectively.

Definition A.8 Let R be a redex of M , assume $\text{occ}(M, \pi) = R$. The function $\text{adr} : \mathcal{T} \times \text{Paths} \times \text{Paths} \rightarrow \mathcal{P}(\text{Paths})$ is defined in a way such that its value $\text{adr}(M, \pi, \xi)$ gives the set of addresses of the residuals of the subterm $\text{occ}(M, \xi)$ after reducing in M with the redex R .

- $R = (\lambda x P Q)$.
 - $\text{adr}(M, \pi, \xi) = \{\xi\}$ if $\pi \approx \xi$,
 - $\text{adr}(M, \pi, \xi) = \{\pi \# \zeta\}$ if $\xi = [\pi :: l :: \lambda x] \# \zeta$,
 - $\text{adr}(M, \pi, \xi) = \{\pi \# \varepsilon \# \zeta \mid \text{occ}(P, \varepsilon) = x\}$ if $\xi = [\pi :: r] \# \zeta$,
 - $\text{adr}(M, \pi, \xi) = \{\xi\}$ if $\xi < \pi$.
- $R = (\mu \alpha P Q)$.
 - $\text{adr}(M, \pi, \xi) = \{\xi\}$ if $\pi \approx \xi$,
 - $\text{adr}(M, \pi, \xi) = \{[\pi :: \mu \alpha] \# l(\text{occ}(M, \pi'), \alpha, \zeta)\}$, if $\xi = \pi' \# \zeta$ with $\pi' = [\pi :: l :: \mu \alpha]$,

- $\text{adr}(M, \pi, \xi) = \{[\pi :: \mu\alpha] \# l(\text{occ}(M, \pi'), \alpha, \epsilon) \# [r_\alpha :: r :: \zeta] \mid \text{occ}(P, \epsilon) = (\alpha V), \text{ for some } (\alpha V) \leq P\} \text{ if } \xi = [\pi :: r] \# \zeta \text{ and } \pi' = [\pi :: l :: \mu\alpha],$
- $\text{adr}(M, \pi, \xi) = \{\pi\} \text{ if } \xi = [\pi :: l],$
- $\text{adr}(M, \pi, \xi) = \{\xi\} \text{ if } \xi < \pi.$
- $R = (\alpha \mu\beta P).$
 - $\text{adr}(M, \pi, \xi) = \{\xi\} \text{ if } \pi \approx \xi,$
 - $\text{adr}(M, \pi, \xi) = \{[\pi] \# \zeta'\} \text{ if } \xi = [\pi :: r_\alpha :: \mu\beta] \# \zeta, \text{ where } \zeta' \text{ is obtained from } \zeta \text{ by exchanging every occurrence of } \beta \text{ in } \zeta \text{ for } \alpha,$
 - $\text{adr}(M, \pi, \xi) = \{\xi\} \text{ if } \xi < \pi.$
- $R = \mu\alpha(\alpha P).$
 - $\text{adr}(M, \pi, \xi) = \{\xi\} \text{ if } \xi \approx \pi,$
 - $\text{adr}(M, \pi, \xi) = \{\pi \# \xi'\} \text{ if } \xi = [\pi :: \mu\alpha :: r_\alpha] \# \xi',$
 - $\text{adr}(M, \pi, \xi) = \{\xi\} \text{ if } \xi < \pi.$

To avoid capture of free variables, the usual way in the literature is the variable convention of Barendregt ([1]). By identifying terms modulo renaming of bound variables, we may assume that, if M_1, \dots, M_n are terms occurring in a certain mathematical context, the names of the bound and free variables are chosen to be different in all these terms. With these stipulation the above definitions for addresses of residuals remain true.

Example A.9 Let $M = (x (\mu\alpha(\alpha (\lambda xy (\alpha x))) y))$ and $M' = (x \mu\alpha(\alpha ((\lambda xy (\alpha(x y))) y)))$, then $M \rightarrow_\mu M'$. Let us compute according to Definition A.8 some of the values $\text{adr}(M, \pi, \xi)$, where $\pi = [r]$, i.e. we compute the indexes of the subterms of M' .

- If $\xi < \pi$, then the only choice is $\xi = []$ and $\text{adr}(M, [r], []) = \{[]\}.$
- If $\xi = [r, l, \mu\alpha] \# [r_\alpha, l, \lambda x]$, then $\text{adr}(M, \pi, \xi)$ is the first occurrence of y in M . Using the notation of the above definition, we have $\xi = \pi' \# \zeta$, where $\pi' = [r, l, \mu\alpha]$ and $\zeta = [r_\alpha, l, \lambda x]$. Then $l(\text{occ}(M, \pi'), \alpha, \zeta) = [r_\alpha, l, l, \lambda x]$ and $\text{adr}(M, \pi, \xi) = \{[r, \mu\alpha] \# [r_\alpha, l, l, \lambda x]\}.$
- If $\xi = [r, r]$, then $\text{adr}(M, \pi, \xi)$ is the second occurrence of y in M . Using the notation of the above definition, we obtain $\xi = [\pi :: r] \# \zeta$, and $\zeta = []$. We have $\text{occ}(P, \epsilon) = \{[], [r_\alpha, r]\}$ as the occurrences of (αV) in $(\alpha (\lambda xy (\alpha x)))$. Then $\text{adr}(M, \pi, \xi) = \{[r, \mu\alpha] \# [] \# [r_\alpha, r], [r, \mu\alpha] \# [r_\alpha, l, r] \# [r_\alpha, r]\}.$

Definition A.10 We define the function $\text{des}_0 : \mathcal{T} \times \text{Paths} \times \text{Paths} \rightarrow \mathcal{P}(\mathcal{T})$ as follows. Let $M \rightarrow^R N$ and $\text{occ}(M, \pi) = R$. The value $\text{des}_0(M, \pi, \xi)$ determines the set of descendants of the subterm of M at address ξ after reducing in M with the redex at address π . That is, $\text{des}_0(M, \pi, \xi) = \{\text{occ}(N, \zeta) \mid \zeta \in \text{adr}(M, \pi, \xi)\}.$ The above function is not everywhere defined. More exactly, for every address π there are terms which has no descendants w.r.t. the reduction with the redex R at address π .

- If $R = (\lambda x P Q)$, then the terms which have no residuals are those with addresses π and $[\pi :: l]$, namely R itself and $\lambda x P$.
- If $R = (\mu\alpha P Q)$, then R only has no residual.
- If $R = (\alpha \mu\beta P)$, then R and $\mu\beta P$, that is, the terms with addresses π and $[\pi :: \alpha]$, have no residuals.

- If $R = \mu\alpha(\alpha P)$, then $\mu\alpha(\alpha P)$ and αP , namely, the terms with addresses π and $[\pi :: \mu\alpha]$, have no residuals.

Example A.11 Returning to Example A.9, the values for des_0 give the terms corresponding to the addresses provided by adr .

- If $\xi < \pi$, then $des_0(M, [r], []) = \{occ(M', [])\} = \{M'\}$.
- If $\xi = [r, l, \mu\alpha] \# [r_\alpha, l, \lambda x]$, then $des_0(M, \pi, \xi) = \{y\}$.
- If $\xi = [r, r]$, then $des_0(M, \pi, \xi) = \{y\}$.

We describe now the set of residuals of a given term with respect to a reduction sequence. To this end, we determine the set of addresses for the residuals. The definition below is an extension of Definition A.8.

Definition A.12 Let σ be a sequence of addresses. We define the values $adr(M, \sigma, \xi)$ of the function $adr : \mathcal{T} \times Paths^{<\infty} \times Paths \rightarrow \mathcal{P}(Paths)$ by induction on $|\sigma|$.

- $adr(M, [\pi], \xi) = adr(M, \pi, \xi)$, as in Definition A.8, if $\sigma = [\pi]$.
- $adr(M, [\pi :: \sigma_1], \xi) = \bigcup \{adr(N, \sigma_1, \zeta) \mid \zeta \in adr(M, \pi, \xi)\}$, where $M \rightarrow^R N$, in other words, $N = des(M, \pi, [])$ and $occ(M, \pi) = R$.

Example A.13 Let $M = (x (\mu\alpha(\alpha (\lambda v((u v) v) (\alpha x))) y))$ and $M' = (x \mu\alpha(\alpha ((\lambda v((u v) v) (\alpha(x y))) y)))$, where $M \rightarrow M'$.

We calculate $adr(M, [[r], [r, \mu\alpha, r_\alpha, l]], [r, l, \mu\alpha, r_\alpha, r, r_\alpha])$, that is, we are concerned with the addresses of the residuals of x in (αx) when implementing the reduction steps with $(\mu\alpha(\alpha (\lambda v((u v) v) (\alpha x))) y)$ and $(\lambda v((u v) v) (\alpha(x y)))$. We have $adr(M, [[r], [r, \mu\alpha, r_\alpha, l]], [r, l, \mu\alpha, r_\alpha, r, r_\alpha]) = \{adr(M', [[r, \mu\alpha, r_\alpha, l]], [r, \mu\alpha, r_\alpha, l, r, r_\alpha, l])\}$ and $adr(M', [[r, \mu\alpha, r_\alpha, l]], [r, \mu\alpha, r_\alpha, l, r, r_\alpha, l]) = \{occ(M', [r, \mu\alpha, r_\alpha, l, l, r, r_\alpha, l])\}$, $occ(M', [r, \mu\alpha, r_\alpha, l, l, r, r_\alpha, l])$.

With this in hand, we are able to define the set of descendants as a set of subterms.

Definition A.14 Let σ be a sequence of addresses. We define the values $des(M, \sigma, \xi)$ of the function $des : \mathcal{T} \times Paths^{<\infty} \times Paths \rightarrow \mathcal{P}(\mathcal{T})$ by $des(M, \sigma, \xi) = \{occ(M, \zeta) \mid \zeta \in adr(M, \sigma, \xi)\}$.

Example A.15 Continuing the previous example, $des(M, \sigma, \xi) = \{occ(M, \zeta) \mid \zeta \in adr(M, \sigma, \xi)\} = \{x\}$, where $M = (x (\mu\alpha(\alpha (\lambda v((u v) v) (\alpha x))) y))$ and we intend to find the descendants of x at address $\xi = [r, l, \mu\alpha, r_\alpha, r, r_\alpha]$ with respect to the redexes at addresses $\sigma = [[r], [r, \mu\alpha, r_\alpha, l]]$.

Definition A.16 Let $M_1 \rightarrow^{R_1} M_2 \rightarrow^{R_2} \dots \rightarrow^{R_n} M_{n+1}$. Suppose $occ(M_1, \pi) = R$ is a redex of M_1 . Then R is involved in σ , if there is a $\sigma_i = [R_1, \dots, R_i]$ ($1 \leq i < n$) such that $R_{i+1} \in des(M_1, \sigma_i, \pi)$.

Example A.17 Let $M = (\lambda x(x x) (\lambda y y y))$ and $R = occ(M, [r]) = (\lambda y y y)$. Assume $\sigma = [[], [r]]$, that is, $M \rightarrow_\beta ((\lambda y y y) (\lambda y y y)) \rightarrow_\beta ((\lambda y y y) y)$. Then R is involved in σ , since one of its descendants is a member of σ .

We need to clarify one more notion defined informally in Definition 3.13.

- Definition A.18**
1. Let $\text{occ}(M, \xi) = N$ for some $N \leq M$. We say that N has n arguments, if $(N M_1 \dots M_n)$ is the maximal subterm of M of this form for the underlying occurrence. N has 0 arguments, if N is not the left hand side of an application. Observe that the definition is equivalent to saying that n is the length of the end segment of ξ consisting entirely of l -s.
 2. Let $\text{occ}(M, \pi) = x$, let us denote by $\text{arg}(M, \pi)$ the number of arguments of the occurrences of x in M . We write $\text{sumarg}(M, x) = \sum \{\text{arg}(M, \pi) \mid \text{occ}(M, \pi) = x\}$, if x has at least one occurrence in M , otherwise let $\text{sumarg}(M, x) = 0$.